NUCLEAR AND CYTOPLASMIC FUNCTIONS OF A YEAST HETEROGENEOUS NUCLEAR RIBONUCLEOPROTEIN

By

ANNETTE MARIE KRECIC

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

1998

This work is dedicated in memory of my grandfather Walter Reed Oberdorf and Burke Patrick Derr.

ACKNOWLEDGEMENTS

With praise to Him in Whom all things are possible, I give thanks for being graced with a loving and supportive circle of family and friends. I also thank my parents, who have believed in me with enduring faith in my ability to achieve my goals. They have nurtured me both emotionally and spiritually, and they continue to be a source of strength and inspiration for daily living.

I would also like to express my appreciation to my mentor, Dr. Maurice S. Swanson, whose passion for science and commitment to excellence have been a constant source of motivation. His critical and creative thinking skills are ones that I strive to emulate. I also acknowledge my committee members, Alfred S. Lewin, Ph.D., Sue A. Moyer, Ph.D., William W. Hauswirth, Ph.D., and Thomas W. O'Brien, Ph.D., for their guidance and commitment throughout my graduate training. I further acknowledge the entire MGM faculty for their perennial "open door" policy, which has provided a welcoming atmosphere in which to do science. I also extend a warm thank you to Joyce Connors, Bradley Moore, and other members of the staff of the Department of Molecular Genetics and Microbiology for their assistance throughout my graduate education, especially in meeting all of the requirements for this degree.

Science thrives on competition and collaboration, and I could not have asked for a better group of people to work with than the members of the Swanson lab, past and present. They have been challenging and cooperative, depending on what the situation

required. Above all, they have been my friends and have always helped to smooth out the bumps along the way, both personally and professionally. For all this, I am sincerely thankful. Without them, I would not have been able to "make it so." I offer a special thank you to Priscilla Non, whose assistance has been one of the few, but valuable, constants during my research, and to my former classmate Penni Black, whose friendship is among the best rewards of my graduate school experience. I also wish to acknowledge my collaborators, Lionel Minvielle-Sebastia and Walter Keller, for their important contributions during the final stages of my research.

Finally, I would like to thank my husband Mike for his endless patience, compassion and support. He has made it easier to continue on this journey, and for that I am deeply and humbly grateful.

TABLE OF CONTENTS

Abstract of Dissertation Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy

NUCLEAR AND CYTOPLASMIC FUNCTIONS OF A YEAST HETEROGENEOUS NUCLEAR RIBONUCLEOPROTEIN

By

Annette Marie Krecic

May 1998

Chairman: Dr. Maurice S. Swanson
Major Department: Molecular Genetics and Microbiology

Nascent pre-mRNAs are bound by many proteins including the heterogeneous nuclear ribonucleoproteins (hnRNPs), a family of abundant nuclear proteins that exhibit sequence-preference binding to RNA. It has been postulated that hnRNPs bind pre-mRNAs to generate transcript-specific ribonucleoprotein complexes that are the substrates for subsequent pre-mRNA processing and nucleocytoplasmic mRNA transport. HnRNPs may function to ensure the fidelity of gene expression by promoting the formation of RNA structures conducive to pre-mRNA processing events and by recruiting trans-acting factors required for pre-mRNA processing and mRNA transport.

This study focuses on a novel yeast hnRNP related to a subfamily of metazoan hnRNPs that are concentration-dependent regulators of splice-site selection during pre-mRNA splicing. Nab4p is essential for cell viability, but unlike other yeast hnRNPs, the intracellular concentration of Nab4p is crucial for normal cell growth. A twofold increase

in Nab4p inhibited cell growth and continued overexpression resulted in nuclear accumulation of poly(A)$^+$ RNA, accumulation of aberrantly long transcripts and cell death. Nab4p interacts with Nab2p, which influences both polyadenylation and mRNA export, suggesting that Nab4p may be involved in these processes. However, recessive *nab4* mutants only showed defects in 3'-end processing.

In contrast to recent evidence that Nab4p is essential for 3'-end cleavage and polyadenylation, Nab4p was required only to regulate cleavage site selection. In the absence of Nab4p, purified cleavage factors utilized discrete *CYC1* precursor RNA cryptic cleavage sites sequentially in a 3'→5' direction. The addition of Nab4p prevented cleavage at these alternative sites in a concentration-dependent fashion. This is the first evidence that the endonuclease complex may cleave in a processive manner, and suggests that Nab4p restricts cleavage to specific sites. These results indicate that a conserved function of hnRNPs is to regulate endonucleolytic cleavage site selection during pre-mRNA processing.

Many hnRNPs shuttle between the nucleus and cytoplasm, including Nab4p. During diauxic shift from fermentable to non-fermentable carbon sources, Nab4p accumulated in the cytoplasm of cells with a concomitant decrease in the nucleus and no change in the steady-state level of Nab4 protein. These observations were specific to Nab4p among the shuttling yeast hnRNPs, strongly suggesting unique cytoplasmic functions for some hnRNPs.

INTRODUCTION

Overview of Eukaryotic Pre-mRNA Processing and Nucleocytoplasmic Transport of mRNA

Normal cell growth requires the proper execution of gene expression. In eukaryotic organisms, genes are transcribed by one of three RNA polymerases, each responsible for the production of specific classes of RNA molecules. These precursor RNAs undergo a series of post-transcriptional modifications that are also specific to different classes of RNAs. The majority of these modifications occur in the nucleus of cells and include alterations to the 5' and 3' ends of the RNA molecule, removal of unexpressed or otherwise dispensable sequences, and specific base modifications. Most mature RNAs, including messenger RNA (mRNA), ribosomal RNA (rRNA) and transfer RNA (tRNA), are ultimately transported from the nucleus to the cytoplasm where their functions are coordinated to synthesize proteins. Exceptions to this are the small nuclear RNAs and small nucleolar RNAs.

Following transcription initiation by RNA polymerase II (pol II), pre-mRNAs undergo a series of maturation steps giving rise to mRNAs which are competent for transport to the cytoplasm and subsequent translation into protein. These post-transcriptional steps include 5'-end capping, pre-mRNA splicing, and 3'-end formation. These mature mRNAs are then competent for nucleocytoplasmic transport. The events described here as post-transcriptional are those that occur after the initiation of transcription. These processes provide multiple levels for cells to regulate gene

expression, and will be reviewed. Transcription initiation, elongation and termination are also important regulatory mechanisms of gene expression. However, these processes will not be discussed in detail.

5'-End Processing: Cap Formation

Pre-mRNA and mRNA have at their 5' ends a complex methylated structure called a cap. This structure is formed via modifications at the terminal nucleotide including the addition of a guanylate residue via a 5'-5' linkage and subsequent methylation (Shatkin, 1976). In addition, no nascent cellular pol II transcripts greater than 20-30 nucleotides have been detected without a cap, indicating that capping must occur soon after transcription initiation (Coppola et al., 1983). It was originally thought that the cap simply aids in translation initiation in the cytoplasm. However, more recent studies have suggested that the m^7G cap is necessary for the export of U snRNAs and has a stimulatory role in pre-mRNA splicing, mRNA 3'-end formation and mRNA export (reviewed in Lewis and Izzauralde, 1997). These stimulatory effects appear to be mediated by a primarily nuclear cap-binding protein complex. For example, in the dipteran *Chironomous tentans*, the cap-binding protein CBP20 accompanies the large Balbiani ring mRNA into the cytoplasm of cells, suggesting that CBP20 contributes to the stimulatory effect of the cap structure on mRNA export. During export, this large mRNA is exported in a 5'→3' direction, leading to the further suggestion that the cap structure and its associated proteins may be involved in orienting mRNAs for recognition by the nuclear pore complex (NPC) during mRNA export. Finally, it was recently shown by Shatkin and colleagues that capping and transcription elongation are coupled by RNA pol II, which could explain the selective capping of pol II transcripts (Yue et al., 1997).

Internal Sequence Modification: Pre-mRNA Splicing

Most higher eukaryotic pre-mRNAs contain expressed sequences, or exons, and intervening sequences, or introns, which generally do not contain protein-coding information (Kramer, 1995). These exons are relatively short, ranging from 10-400 nucleotides, while introns, numbering anywhere from one to >50 in a single pre-mRNA, can be much longer, in some cases extending up to 200,000 nucleotides. In striking contrast, only 2-5% of all yeast nuclear-encoded genes contain intron sequences and most of these contain just one intron, usually close to the 5' end of the pre-mRNA and generally less than 400 nucleotides in length (Rymond and Rosbash, 1992). Yeast exons vary in size ranging from a few nucleotides to several thousand nucleotides. The removal of introns is essential for gene expression, as they often contain in-frame stop codons that lead to truncated non-functional proteins, or even toxic protein products, upon translation (Kramer, 1995). Despite their necessary removal, introns can also provide a means of regulating gene expression. Alternative splicing pathways which leave an intron intact can down-regulate gene expression. In addition, alternative removal and usage of introns and exons can produce multiple isoforms of a protein from the same pre-mRNA, thus extending the protein-coding capacity of the genome. While there are only two documented cases of regulated splicing in *Saccharomyces cerevisiae* (Eng and Warner, 1991; Engebrecht et al., 1991), alternative splicing has not been observed in yeast (Rymond and Rosbash, 1992).

Through many years of both *in vivo* and *in vitro* studies in a variety of eukaryotic organisms from yeast to humans, the biochemistry of splicing has been well established and occurs via a two step process (reviewed in Kramer, 1995). First, an endonucleolytic

cleavage reaction separates the intron from the 5' exon. Next, the 3' end of the intron is cleaved, concurrent with ligation of the 5' and 3' exons. Despite the exceedingly large size of some introns, their correct removal requires limited, highly conserved sequence elements located near the junctions between introns and exons. However, these cis-elements and the trans-acting factors that recognize them are critical to precisely execute the splicing reaction. Nature has shown us that deviation by even a single nucleotide can shift the sites of endonucleolytic cleavage and ligation, spelling disaster for cellular metabolism. Such deviations have been discovered as the molecular cause of many genetic diseases, including β-thalassemias and cystic fibrosis (Krawczak et al., 1992; Antoniou, 1995).

3'-End Processing: Cleavage and Polyadenylation

Except for mRNAs encoding the replication-dependent histones in metazoan cells, all eukaryotic mRNAs have at their 3' end a polyadenylate tail (reviewed in Keller, 1995; Colgan and Manley, 1997). While the best evidence for the function of the poly(A) tail is in the efficiency of mRNA translation initiation and mRNA stability, it has been proposed that the poly(A) tail can influence almost all aspects of mRNA biogenesis. More recently, it has emerged that polyadenylation is also important for splicing, transcription termination and mRNA export, which may provide additional means by which cells regulate gene expression. The details of how these processes might be linked to polyadenylation are largely unknown.

While the formation of pre-mRNA 3' ends is an important event during mRNA biogenesis in all eukaryotes, the biochemistry of 3'-end formation has been studied most extensively in mammalian cells and the *S. cerevisiae*. Through the combination of *in*

vivo and *in vitro* studies, a general mechanism has emerged in which poly(A) tails are formed during a two-step process that begins with an endonucleolytic cleavage near the 3' end of the pre-mRNA, followed by the addition of adenosine residues to the upstream cleavage product (reviewed in Wahle and Keller, 1992, 1996). However, the details of this mechanism and the regulation of this process appear to vary somewhat among eukaryotes. For example, the cleavage and polyadenylation reactions are tightly coupled in mammalian cells, and the maximal length of poly(A) tails is species-specific, averaging 200-250 nucleotides in mammalian cells. In yeast, the two reactions of 3'-end formation are not as tightly coupled, and the maximal length of poly(A) tails is 50-70 nucleotides. Like splicing, the execution of 3'-end formation requires both cis-elements in the transcript as well as trans-acting factors. While not as well characterized as the splicing reaction, the mechanism of 3'-end formation clearly involves recognition and binding by these trans-acting factors to sequences near the site of endonucleolytic cleavage. Likewise, the activities of the trans-acting factors are coordinated by the formation of a multi-subunit enzymatic complex within which 3'-end formation is executed. With the exception of the almost invariant poly(A) signal in mammalian pre-mRNAs, the RNA sequences required for 3'-end formation are not as conserved in eukaryotes as those required for splicing. The relative position and close spacing of these sequences is important, unlike 5' and 3' splice sites, which can be separated by great lengths of intervening sequences. Alternative usage of 3'-end endonucleolytic cleavage sites occurs in eukaryotes (Edwalds-Gilbert et al., 1997), often involving 5'-splice site-related or intron sequences (Lutz and Alwine, 1994; Furth et al., 1994; Lou et al., 1996). In one instance, the relative abundance of a basal cleavage factor appears to be important

(Takagaki et al., 1996). However, the mechanisms regulating cleavage site selection remain speculative.

Much of our current understanding of pre-mRNA 3'-end processing has come from the study of *in vitro* reconstitution of 3'-end formation using biochemically purified factors, again primarily using mammalian cells and *S. cerevisiae* (reviewed in Wahle and Keller, 1996, Keller and Minvielle-Sebastia, 1997, and Colgan and Manley, 1997). In both systems, RNA elements have been identified that specify the site of polyadenylation, but these are somewhat different between the two systems. In mammalian pre-mRNAs, these sequence elements flank the site of endonucleolytic cleavage, with most pre-mRNAs containing the hexanucleotide AAUAAA poly(A) signal located upstream and a GU-rich sequence found downstream. Not only the sequence of these elements but also the spacing between them are important in specifying cleavage (MacDonald et al, 1994; Chen et al., 1995), with a preference for cleavage after a CA dinucleotide within the permissive region (Chen et al., 1995). Similar sequences have been identified in yeast, but there are notable differences. In yeast, both sequences lie upstream of the cleavage site, presumably because the relatively short distances between genes in the yeast genome cannot accommodate downstream elements (Manley and Takagaki, 1996). These sequence elements are also degenerate and redundant compared to mammalian poly(A) signals (Guo and Sherman, 1996). While AAUAAA can function as the more proximal positioning element in yeast (Guo and Sherman, 1995), more often this element is defined by an A-rich sequence. The more distal efficiency element is generally AU-rich and can consist of multiple copies of this type of sequence, which have an additive effect

on the efficiency of cleavage (Guo et al., 1995). The site of cleavage in yeast pre-mRNAs often follows the sequence $U/C(A)_n$.

Because of the seemingly unrelated polyadenylation signals found in mammalian pre-mRNAs versus yeast pre-mRNAs, it was generally viewed that the protein machinery specifying 3'-end formation would also be different. In both mammalian and yeast systems, a number of protein factors are required for 3'-end processing, most of which are composed of multiple subunits (reviewed in Wahle and Keller, 1996, Keller and Minvielle-Sebastia, 1997, and Colgan and Manley, 1997). By combining the availability of genetic techniques and the complete genomic sequence of *S. cerevisiae*, the identification of these factors has been especially rapid in yeast. While none of the mammalian factors function in combination with the yeast factors, or vice versa, many of these proteins, quite unexpectedly, appear to have been conserved between *S. cerevisiae* and human cells. In fact, structural homologues have been identified in a variety of eukaryotic organisms, including *Schizosaccharomyces pombe*, *Arabidopsis thaliana*, *Caenorhabditis elegans*, *Drosophila*, *Xenopus laevis* and chicken. However, the extent to which these proteins are functional homologous to their mammalian or *S. cerevisiae* counterparts remains to be determined. It is important to note that at least some polyadenylation signals from the evolutionarily divergent yeasts *S. pombe and S. cerevisiae* are functionally interchangeable, suggesting that the mechanism of pre-mRNA 3'-end formation may be conserved in lower eukaryotes (Humphrey et al., 1991).

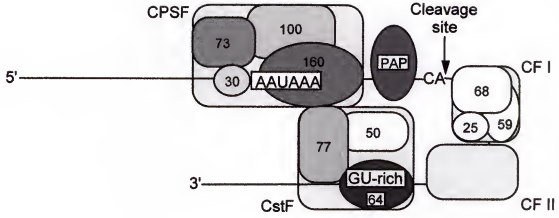
The components of mammalian 3'-end processing are cleavage stimulation factor (CstF) and cleavage factors I_m and II_m (CF I_m and CF II_m), all of which are required for cleavage only, and cleavage and polyadenylation stimulation factor (CPSF) and poly(A)

polymerase (PAP), each required for both cleavage and polyadenylation. In addition, the nuclear poly(A)-binding protein PAB II is involved in elongation of the poly(A) tail and control of poly(A) tail length. Specific cleavage during yeast 3'-end formation requires cleavage factors I and II (CF I and CF II, not to be confused with the similarly named mammalian factors), whereas specific polyadenylation of a pre-cleaved RNA requires CF I in addition to polyadenylation factor (PF I) and poly(A) polymerase (Pap1p). The major cytoplasmic poly(A) tail-binding protein Pab1p has been shown to affect poly(A) tail length both *in vivo* (Sachs and Davis, 1989) and *in vitro* (Minvielle-Sebastia et al., 1997; Amrani et al., 1997b; Kessler et al., 1997), but unlike PAB II in mammalian polyadenylation, Pab1p is not required for poly(A) tail elongation in yeast.

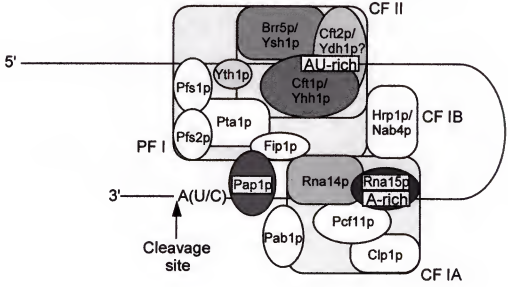
While the molecular mechanisms of mammalian and yeast 3'-end formation are just beginning to emerge, the current models for each share the same general features (reviewed in Wahle and Keller, 1996, Keller and Minvielle-Sebastia, 1997, and Colgan and Manley, 1997). It is clear that a complex array of RNA-protein interactions and protein-protein interactions is important, including both inter-subunit and intra-subunit interactions. Illustrating this is the current mechanistic model of mammalian 3'-end processing (Figure 1A). The AAUAAA sequence in mammalian pre-mRNAs is recognized and bound by CPSF. While there are no homologies to known RNA-binding motifs, the largest subunit of CPSF can be crosslinked to RNA, albeit with limited preference for AAUAAA. Still, it is believed that this RNA-protein interaction mediates the binding of CPSF to the poly(A) signal. Similarly, CstF binds to the GU-rich element, perhaps via the 64 kDa subunit, which contains a typical RBD motif and can be crosslinked to the downstream element. Moreover, the individual binding of either CPSF

Figure 1. Schematic representation of the cleavage and polyadenylation complexes required for mammalian (A) and yeast (B) pre-mRNA 3'-end processing. The large lightly shaded rounded boxes delineate the biochemically purified multi-subunit factors and individual proteins are shown as smaller rounded rectangles or ovals. Structural orthologues are depicted with similar shading. The cis elements required for processing specificity are highlighted in white rectangles on generic 3'-end processing substrates. With the exception of CstF, the relative arrangement of proteins within factors is speculative.

A. Mammalian



B. Yeast



or CstF to RNA substrates is unstable and the increased stability of a complex containing both protein factors indicates that they bind cooperatively. Protein-protein contacts between the largest subunits of CPSF and CstF have been identified and are likely to mediate their cooperative binding. Although the precise interactions of CF I_m and CF II_m remain unclear, all three subunits of CF I_m preferentially bind to RNA substrates containing 3'-end processing signals, and CF I_m further stabilizes the CPSF-RNA complex, suggesting that protein-protein interactions between CF I_m and CPSF are important. Upon recruitment of PAP by CPSF, this assemblage of RNA and proteins specifies the cleavage site and cleaves the pre-mRNA. The source of the endonuclease activity within this cleavage apparatus has not been identified, but it was recently demonstrated that the *Drosophila* protein clipper, a homologue of 30 kDa subunit of CPSF, possesses N-terminal endoribonucleolytic activity (Bai and Tolias, 1998). Following cleavage, CstF, CF I_m and CF II_m are dispensable and polyadenylation ensues, depending only on CPSF and PAP. The first phase of poly(A) addition is distributive, occurring rather slowly until the poly(A) tail is 10-12 residues in length. At this point, PAB II binds to this short tail and stimulates the processive and rapid elongation of the poly(A) tail. Additional molecules of PAB II continue to associate with the growing poly(A) tail, which eventually results in the cessation of the processive reaction. Poly(A) tails are limited to 250-300 nucleotides *in vitro*, corresponding to the average maximal length of poly(A) tails seen *in vivo*.

The poly(A) polymerase, Pap1p, was the first 3'-end processing factor identified in yeast (Lingner et al., 1991). Prior to 1996, only three additional subunits of the yeast polyadenylation machinery had been identified. Two subunits of CF I, Rna14p and

Rna15p, were originally identified genetically, as mutations in *RNA14* and *RNA15* caused abnormal mRNA decay rates and shortening of poly(A) tails *in vivo* (Minvielle-Sebastia et al., 1991). It was later demonstrated that these proteins are interacting components of CF I, required for both cleavage and polyadenylation (Minvielle-Sebastia et al., 1994; Kessler et al., 1996). One subunit of PF I, Fip1p, was isolated via a two-hybrid screen to identify factors that interact with the Pap1p (Preker et al., 1995). Since Rna15p was the only known component of CF I to contain a conserved RNA-binding motif, it was proposed that Fip1p, which also interacts with Rna14p, might tether PF I and the poly(A) polymerase to the CF I-RNA complex. This could stabilize the complex required for polyadenylation and explain the nature of the dual requirement of CF I during 3'-end formation.

In the past two years, a few groups have systematically identified nearly all of the remaining components required for 3'-end processing in yeast (Kessler et al., 1996, 1997; Amrani et al., 1997a; Zhao et al., 1997; Minvielle-Sebastia et al., 1997; Preker et al., 1997). CF I has been separated into two components, CF IA, which contains the gene products of *PCF11* and *CLP1* in addition to *RNA14* and *RNA15*, and the single-subunit CF IB, identified as Hrp1p (Kessler et al., 1997). The composition of CF II and PF I appears to be overlapping, with most of the subunits of CF II (Zhao et al., 1997) also reported as subunits of PF I (Preker et al., 1997).

The independent isolation and identification of some of these proteins has allowed the characterization of their individual effects on 3'-end formation. Ysh1p/Brr5p and Cft1p, identified as a subunits of both CF II and PF I, were first identified as 3'-end processing factors by virtue of homology to subunits of mammalian CPSF (Chanfreau et

al., 1996; Jenny et al., 1996; Stumpf and Domdey, 1996). These reports demonstrated that Ysh1p/Brr5p and Cft1 affect both cleavage and polyadenylation steps. In the case of Cft1p, restoration of cleavage activity upon immunodepletion of Cft1p was possible with CF II alone (Stumpf and Domdey, 1996). However, polyadenylation activity was not restored by the addition of CF II and in fact, reconstitution required addition of both PF I and poly(A) polymerase. These results suggest that Cft1p or another subunit of CF II interact with PF I and perhaps also with Pap1p.

As in the mammalian system, RNA-protein and protein-protein interactions are necessary to coordinate the activities of the yeast 3'-end processing factors (Figure 1B). This aspect has not been extensively examined in the yeast system, but with the identification of the yeast cleavage and polyadenylation factors almost complete, these studies are sure to follow soon. As described previously, CF IA probably binds to RNA via Rna15p. No RNA-binding motifs have been identified in any of the subunits of CF II, but the homology between Cft1p and the AAUAAA-binding largest subunit of CPSF make Cft1p a possible candidate for RNA binding, even though the AAUAAA element is not conserved in yeast. While protein-protein and genetic interactions have been demonstrated between various subunits of CF IA, interactions among CF II subunits have not been described. The same is true of PF I, but it is likely that the various homologues of CPSF subunits found in PF I (Yhh1p/Cft1p, Ydh1p, Ysh1p/Brr5p and Yth1p) interact with one another as do their mammalian counterparts. Also, the PF I homologues of the mammalian CPSF subunits may mediate RNA binding. Alternatively, all three subunits of CF I_m can be crosslinked to substrate RNAs (Rüegsegger et al., 1996), so if PF I contains any homologues to CF I_m, these may in turn mediate binding to the pre-mRNA.

Another question is where, and in what orientation, the yeast factors bind to the RNA substrate to specify cleavage. The only data directly addressing this question is in regard to CF I. When CF I was first separated into CF IA and CF IB, it was shown that only Rna15p could be crosslinked to pre-cleaved wild type or mutant [$\Delta(UA)_6$] RNA substrate (Kessler et al., 1996). However, it was subsequently reported that recombinant Hrp1p could be crosslinked to RNA substrates in a manner which is dependent on the presence of the $(UA)_6$ polyadenylation efficiency element (Kessler et al., 1997). In fact, the efficiency of crosslinking was even greater using a pre-cleaved RNA substrate, and CF II, but not CF IA, stimulated crosslinking of recombinant Hrp1p to both full-length and pre-cleaved substrates.

In addition to Pab1p, Hrp1p is reported to influence poly(A) tail length (Kessler et al., 1997). In these experiments, a pre-cleaved substrate was polyadenylated in the absence of Hrp1p, albeit less efficiently, and the distribution of the polyadenylated product was longer than normal. Addition of recombinant Hrp1p increased the efficiency of polyadenylation and restored the distribution to a more normal length. Interestingly, mutations in *HRP1* were reported to cause shortening of poly(A) tails *in vivo*, similarly to mutations in *RNA14* and *RNA15* (Minvielle-Sebastia et al., 1991). Yet the absence of Hrp1p *in vitro* resulted in hyperpolyadenylation, in contrast to the results seen using extracts from *rna14*, *rna15* and *pap1* mutant strains, which showed little to no polyadenylation activity (Minvielle-Sebastia et al., 1994). However, Hrp1p was not obligatory for poly(A) addition, so again, the role of Hrp1p in polyadenylation seems very different from that of other polyadenylation factors.

Messenger RNA Transport

Nucleocytoplasmic mRNA export

Most of what is known about RNA export from the nucleus comes from *in vivo* studies entailing the microinjection of *Xenopus* oocytes or the examination of yeast mutants that exhibit nuclear accumulation of poly(A)$^+$ RNA (reviewed by Izaurralde and Mattaj, 1995, and Nakielny and Dreyfuss, 1997). While a number of factors implicated in export have been identified by a variety of biochemical techniques, the functional analysis of export factors is limited by the lack of an assay that faithfully recapitulates mRNA export *in vitro*. The export of mRNA from the nucleus consists of two phases: movement of the transcript from the site of transcription to the nuclear pore complex (NPC) followed by translocation across the nuclear envelope through the NPC.

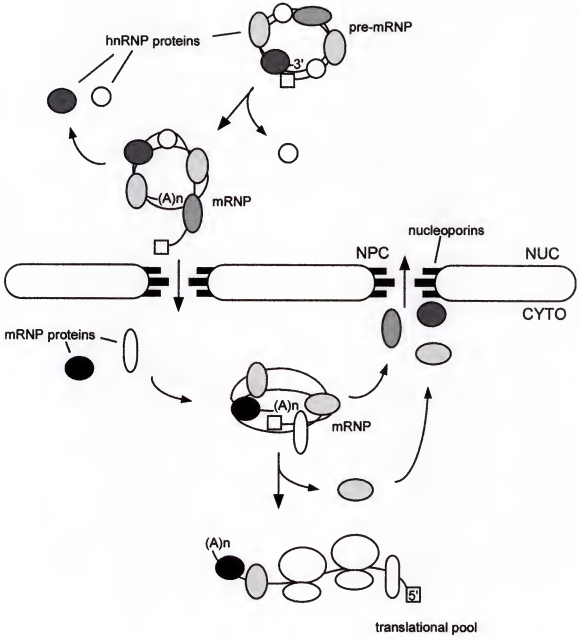
In eukaryotic cells, the nuclear envelope is perforated by NPCs (reviewed in Boelens et al., 1995). These large structures are proteinaceous and consist of an estimated 100 or more polypeptides assembled into a channel, providing the means by which protein and RNA molecules move between the nucleus and cytoplasm. Ions and small molecules are transported through the NPC by passive diffusion, whereas larger molecules greater than ~60 kDa undergo facilitated transport via energy-dependent pathways. It is generally believed that movement through NPCs is bi-directional and that pores are not specialized in terms of the type of RNA or protein substrates that can pass through any given NPC.

The isolation of NPCs from *S. cerevisiae* and the use of genetic screens for yeast mutants that exhibit nuclear accumulation of poly(A)$^+$ RNA has led to the identification of a number of NPC proteins, or nucleoporins (Izzaurralde and Mattaj, 1995). Genetic

and physical interactions have been demonstrated between nucleoporins, and mutations in the corresponding genes often cause multiple defects in addition to their effect on mRNA export. These can include a block to protein import, clustering of NPCs, structural defects in NPCs and the nuclear envelope, and nucleolar fragmentation. Therefore, it is not always clear whether the effect on mRNA export is direct or a downstream effect of these additional defects. There are certain alleles of some nucleoporin genes, including *NUP159/RAT7*, *NUP133/RAT3*, and *NUP49*, which effect RNA export and not protein import, suggesting that these alleles do not simply disrupt the functional integrity of NPCs. Since these three genes interact genetically, their protein products may associate to form a subregion of the NPC specifically needed for RNA export.

The substrates for RNA export are ribonucleoprotein complexes possibly assembled during transcription (see Figure 2). This has been illustrated perhaps best by studies of the very large Balbiani ring (BR) particle of *Chironomous tentans* (reviewed in Mehlin and Daneholt, 1993). This specialized particle, containing a 35 kb transcript and many associated proteins, can be visualized by electron microscopy. These studies revealed that the BR particle assumes a compact folded structure after transcription, but upon docking at the NPC, proceeds to unravel and is translocated through the NPC as a RNP fibril, in a 5'→3' orientation. In addition, some of the associated proteins dissociate from the particle and remain in the nucleus (Alzhanova-Ericsson et al., 1996) while others remain associated during translocation (Visa et al., 1996a, 1996b), indicating that both the structure and the composition of the RNP undergo dynamic changes during transport. Those proteins remaining associated may interact with the NPC to facilitate

Figure 2. Generalized model for mRNA export. In the nucleus, RNA polymerase II transcripts are bound by hnRNPs during and following transcription to form pre-mRNP complexes. These RNP complexes become the substrates for mRNA export, which occurs in a 5'→3' direction through the nuclear pore complex (NPC). As the mRNA is translocated through the NPC, associated hnRNPs may interact with nucleoporins or dissociate from the mRNP. Other hnRNPs accompany the mRNA to the cytoplasm where they may remain associated or are exchanged for cytoplasmic mRNA binding proteins. mRNPs may then be transported through the cytoplasm and/or enter the translational pool. Circles and ovals represent hnRNP and mRNP proteins, the dark lines within the NPC represent nucleoporins, and the 5' cap structure and cap-binding complex are represented by squares.


hnRNP proteins
pre-mRNP
3'
(A)n
mRNP
nucleoporins
NPC
NUC
CYTO
mRNP proteins
(A)n
mRNP
(A)n
5'
translational pool

translocation of the RNP, perhaps with those nucleoporins that appear to be specifically required for RNA export. At least two proteins in the BR RNP complex, hrp36 and the cap binding protein CBP20, bind to the pre-mRNA during transcription and accompany the particle to the NPC. The hnRNP A/B-type protein hrp36 remains associated with the BR mRNA during translocation through the NPC and is present in polysomes during translation of the BR mRNA.

Since the substrates for mRNA export are RNA-protein complexes, mRNA export has been viewed as a process involving nuclear protein export, the mechanism of which seems to be fundamentally similar to nuclear protein import. Like protein import, RNA export is temperature- and energy-dependent, and involves the hydrolysis of guanosine triphosphate (GTP) by a GTPase, a guanine nucleotide exchange factor (GEF), and a GTPase activating protein (GAP). Many studies have demonstrated that the yeast and vertebrate homologues of the GTPase Ran/TC4, RCC1 (a nuclear RanGEF), and *RNA1* (a cytoplasmic RanGAP) are involved in both nuclear protein import and mRNA export (reviewed by Moore and Blobel, 1994; Izaurralde and Mattaj, 1995; and Sazer, 1996). For example, in the temperature-sensitive hamster cell line tsBN2, loss of RCC1 protein at the non-permissive temperature results in the accumulation of poly(A)$^+$ RNA in the nucleus of these cells. RCC1 is also required for the accumulation of U3 snRNA in the nucleolus, suggesting that RCC1 may be involved in transport of RNA within the nucleus and not simply translocation through the NPC during RNA export. Similarly, mutations in the yeast homologues of RCC1 and RanGAP, *PRP20* and *RNA1*, respectively, lead to an apparent block in mRNA export. However, mutations in these yeast genes also give rise to a variety of pleiotropic effects including defects in protein import, RNA

processing, and alterations in nuclear structure. Therefore, it has been difficult to determine the primary role(s) of the RanGTPase pathway and characterize any direct effect on mRNA export.

A number of Ran-binding proteins have been identified including the mammalian BP1, a protein that interacts with GTP-bound Ran and acts as a co-activator of RanGAP, and BP2/Nup358, a Ran-binding nucleoporin (Sazer, 1996). Although direct binding to the essential yeast Ran homologue Gsp1p has not been reported, the yeast nucleoporin Nup1p contains sequences similar to the core Ran-binding domain of BP1. While these interactions may only reflect those required during protein import, continued identification of proteins that interact with components of the RanGTPase pathway should help clarify the role of these proteins in RNA export.

RNA export is also carrier mediated, with the export of various classes of RNA being specifically saturable (Jarmolowski et al., 1994). This implies that the factors mediating RNA export are, at least in part, class specific. It has also been observed that certain RNA elements influence mRNA export, including the 5' m^7G cap structure, splicing signals and introns, and the 3' poly(A) tail or 3' stem-loop of histone mRNAs (Nakielny and Dreyfuss, 1997). Therefore, it is likely that some of the factors mediating mRNA export interact with these cis-elements. For example, by using various cap derivatives on U1 snRNAs as competitors of export, it was demonstrated that U snRNAs require a saturable factor that binds to the cap structure (Jarmolowski et al., 1994). This led to the identification of the nuclear cap-binding complex that has been implicated in mRNA export as well (Visa et al., 1996b). However, a 5' cap, while able to stimulate mRNA export, appears to be more important in the export of U snRNAs than mRNAs.

Studies of viral RNA export have also provided insight into the mechanisms for cellular RNA export. This has been best described for the export of retroviral transcripts. The export of partially spliced or unspliced HIV transcripts necessary to complete the viral life cycle requires the Rev response element (RRE), a viral RNA sequence, which is bound by the essential RNA-binding-protein Rev. This protein contains a leucine-rich nuclear export signal (NES) that mediates the export of intron-containing viral transcripts. Interestingly, microinjection of excess Rev NES into *Xenopus* oocytes blocks the export of U snRNA and 5S rRNA, but not mRNA, tRNA or ribosomal subunits. Rev functions in *Xenopus* oocytes and yeast, suggesting that Rev-mediated export may exploit an export pathway for cellular RNAs or at least a conserved component of a cellular RNA export pathway. Constitutive transport elements (CTEs) have been identified in other retroviruses, which mediate the transport of unspliced viral transcripts independent of a viral protein (Ernst et al., 1997). These elements can functionally replace Rev and the RRE in the export of HIV intron-containing transcripts and can mediate the export of intron-containing cellular RNAs that would otherwise be retained in the nucleus. A cellular protein that binds to CTEs has not been identified but may act similarly to Rev and provide insight regarding cellular RNA export.

Cellular proteins that interact with Rev have been identified, and include a number of NPC or NPC-associated proteins. One of these, the yeast Rip1 protein, is not essential for yeast cell viability but is important for Rev-mediated RNA export (Stutz et al., 1995). Recently, a yeast protein that interacts with Rip1p, Gle1p, was shown to contain a leucine-rich NES. The deletion of this NES results in the rapid accumulation of

poly(A)$^+$ RNA in the nucleus of cells (Murphy and Wente, 1996), suggesting a role for leucine-rich nuclear export signals in mRNA export.

A number of hnRNPs, both yeast and vertebrate, have also been implicated in mRNA export. The best characterized of these are the hnRNP A/B-type proteins, which contain a nuclear localization sequence, called M9, which has been shown to mediate nuclear import and export of these proteins as well as export of mRNA (reviewed in Nakielny and Dreyfuss, 1997). The M9 sequence interacts with transportin, a protein with similarity to the nuclear protein import factor importin β. As an NLS-receptor, transportin mediates the nuclear import of M9-containing proteins. Since transportin binds M9 to mediate import, it is possible that it may also function as a nuclear export receptor. However, NES receptors have not been identified. In yeast, strains carrying mutant alleles encoding the hnRNPs Npl3p and Nab2p exhibit nuclear accumulation of poly(A)$^+$ RNA during growth under non-permissive conditions (Russell and Tollervey, 1994; Singleton et al., 1995; Lee et al., 1996; Anderson, 1995). A direct role for these proteins in mRNA export has not been established.

Finally, the poly(A) tail and the process of 3'-end formation of pre-mRNAs is important in the regulation of mRNA export. This was first demonstrated by experiments which evaluated the ability of processing signals directing histone 3'-end formation or polyadenylation to promote export of an intronless mRNA (Eckner et al., 1991). Both processing pathways enhanced mRNA export. However, mRNAs with a mature histone 3'-end generated by a cis-acting ribozyme were transport deficient, suggesting that histone 3'-end formation and mRNA export are coupled. It was later suggested that 3'-end cleavage and polyadenylation is also linked to mRNA export, as a long stretch of

poly(A) near the 3'end of a transcript generated by a cis-acting ribozyme was not sufficient to direct mRNA export (Huang and Carmichael, 1996). How these two processes are coupled is unknown.

Cytoplasmic mRNA localization

Transport and sorting of mRNA in the cytoplasm is receiving increasing attention as a means for regulating gene expression in a wide range of organisms (reviewed in Wilhelm and Vale, 1993; St. Johnston, 1995; Bassell and Singer, 1997; and Nasmyth and Jansen, 1997). Cytoplasmic localization of mRNA is believed to provide a more efficient means of sorting cytosolic proteins, by spatially controlling translation to restrict or promote protein synthesis in specific subcellular locales. This has been well characterized in *Drosophila*, where protein gradients are responsible for establishing the body plan during embryogenesis. Although the mechanisms of mRNA localization remain poorly defined, a general pathway has been proposed (Wilhelm and Vale, 1993). As in nuclear mRNA export, the substrates for mRNA localization are believed to be RNP complexes. Therefore, the first step in mRNA localization would be the formation of a RNP particle. This particle is then translocated to the appropriate destination where it becomes anchored to the cytoskeleton and is subsequently translated into protein. The existence of this general pathway has been supported by a battery of studies, which have begun to provide some of the details of the mechanisms involved in mRNA localization.

Since all of the mechanisms involved in the regulation of gene expression described thus far require RNA elements, it is not surprising that mRNA localization also requires sequences within the transcript, which have been exclusively limited to the 3' untranslated region (3'-UTR) of asymmetrically distributed mRNAs (Decker and Parker,

1995). Unlike those RNA elements required for splicing or 3'-end formation, those mRNA localization signals that have been precisely mapped are relatively large, on the order of hundreds of nucleotides.

The existence of cis-elements that direct mRNA localization suggests that RNA-binding proteins recognizing these 3'-UTR sequences, or secondary structures formed within them, will also be important in directing the cytoplasmic trafficking of mRNA. A number of cytoplasmic RNA-binding proteins involved in the localization of specific mRNAs have been identified, the best characterized of which is the *Drosophila* protein staufen (St. Johnston, 1995). Deletions in the *staufen* gene result in mislocalization of maternal mRNAs, and staufen has the ability to bind to both microtubules and the 3'-UTR of at least one maternal mRNA, suggesting that it may be an anchoring protein.

Various aspects of the model for mRNA localization have been supported by studies employing the microinjection of myelin basic protein (MBP) mRNA into living neurons. In situ hybridization revealed the formation of RNA 'particles' or granules that seemed to contain multiple copies of the injected mRNA (Ainger et al., 1993). More recently, the movement of these granules into neuronal processes was examined (Knowles et al., 1996). MBP RNA granules were found to move along microtubules in one direction, at a rate consistent with transport by a motor-driven process. It appears that both microtubule and microfilament networks are utilized for the transport of mRNAs, as treatment of cells with drugs that specifically disrupt these cytoskeletal components can lead to the mislocalization of mRNAs (Nasmyth and Jansen, 1997; Bassell and Singer, 1997). In addition, the roles of a number of microtubule-associated or actin-binding proteins in mRNA localization are currently under investigation.

The importance of mRNA localization is not restricted to multicellular organisms with complex developmental programs. In *S. cerevisiae*, the Ash1 protein is a cell-fate determinant, which is localized asymmetrically between mother and budding daughter cell during cell division (Long et al., 1997; Takizawa et al., 1997). The *ASH1* mRNA is preferentially distributed to the daughter cell. This localization requires the 3'-UTR of the ASH1 mRNA and an intact actin cytoskeleton. This 3'-UTR is necessary and sufficient to localize a heterologous mRNA in the same asymmetric pattern seen for *ASH1* mRNA during cell division and localization did not appear to be dependent upon microtubules or specific to any stage of the cell cycle (Long et al., 1997). Thus, it appears that mRNA localization to sort proteins asymmetrically is utilized in a variety of eukaryotic cell types.

Regulation of the Initiation of Translation

Translation of an mRNA begins with the ordered assembly of a ribosome onto the mRNA (Curtis et al., 1995). First, a complex of proteins, including the cap-binding protein eIF-4E and an RNA helicase binds the 5' cap structure. Next, the 43S small ribosomal-initiation complex scans the 5' untranslated region (5'UTR) until it recognizes an AUG initiation codon in a favorable context. After an AUG has been chosen, the large ribosomal subunit enters the pre-initiation complex and protein synthesis ensues.

Translation is generally regulated by modulating the rate of translational initiation, the accessibility of an mRNA to translation, and the length of the poly(A) tail (Curtis et al., 1995). Sequences within the 5'- and 3'-UTRs of mRNAs are usually involved in these cytoplasmic mechanisms for regulating gene expression. For example, stable secondary structures within the 5' UTR can block ribosome binding or scanning

prior to translational initiation. Alternatively, translational masking can repress translation by sequestering an mRNA into a translationally silent mRNP. Interestingly, translationally silent mRNPs are assembled in *Xenopus* oocytes only when mRNAs are transcribed in the nucleus and not when mRNAs transcribed *in vitro* are injected into the nucleus or cytoplasm (Bouvet and Wolffe, 1994). These observations established a functional connection between transcription and assembly of translationally silent mRNPs.

The regulation of the length of poly(A) tails by polyadenylation and deadenylation is also important to translation, with a longer poly(A) tail increasing the rate of translational initiation (Richter, 1995). In *Xenopus* oocytes, cytoplasmic polyadenylation is directed by specific sequences in the 3'-UTR, including a U-rich cytoplasmic polyadenylation element (CPE) and the nuclear polyadenylation sequence AAUAAA. Both nuclear and cytoplasmic activities are required for cytoplasmic polyadenylation (Fox et al., 1989; Varnum et al., 1992), including RNA-binding activities. The nuclear polyadenylation factors CPSF and PAP recognize CPEs (Bilger et al., 1994), and a cytoplasmic poly(A) polymerase closely related to the nuclear PAP has been identified (Ballantyne et al., 1995). It has been suggested that CPSF binds the CPE in the cytoplasm and may be a core component of the cytoplasmic polyadenylation machinery (Bilger et al., 1994). Two RNA-binding proteins that bind to CPEs have been identified in *Xenopus*. One of these proteins, CPEB, mediates cytoplasmic polyadenylation during *Xenopus* oocyte maturation (Hake and Richter, 1994). The other, ElrA, is similar to the ELAV proteins of *Drosophila*, each of which contains three

conserved RNA-binding domains. ElrA requires an embryonic-type CPE to interact with certain *Xenopus* mRNAs *in vivo* (Wu et al., 1997).

Structures of Pre-mRNA/mRNA-Binding Proteins

It is generally believed that the execution of pre-mRNA processing events is controlled by a multitude of factors including both cis-elements found in the nucleotide sequence of pre-mRNAs as well as trans-acting factors that bind to pre-mRNAs. Pre-mRNAs and mRNAs are found in association with hnRNPs throughout their nuclear residency, and a subset of hnRNPs has the ability to shuttle between the nucleus and cytoplasm. In some cases, these proteins are associated with cytoplasmic mRNA. These associations make hnRNPs likely candidates for regulators of pre-mRNA processing and mRNA nucleocytoplasmic transport during gene expression. As a paradigm, this is illustrated quite elegantly in the case of pre-mRNA splicing (Kramer, 1995). There are conserved nucleotide sequences at the 5' and 3' ends of exons and within introns which are required for their removal. These sequences are in turn recognized and bound by transacting factors that carry out the excision of the intron and ligation of the flanking exons. These reactions occur within a large macromolecular structure, the spliceosome, which consists of several RNA-protein subunits, or ribonucleoprotein (RNP) complexes. Through the ordered assembly and disassembly of the spliceosome, these small nuclear RNPs, or snRNPs, facilitate pre-mRNA splicing with incredible fidelity. In addition to the essential role of snRNPs in pre-mRNA splicing, additional RNA-binding proteins can influence the selection of splice sites, including certain hnRNPs.

The family of pre-mRNA/mRNA-binding proteins known as hnRNPs has been studied extensively in a variety of metazoan cells (Dreyfuss et al., 1993; Swanson, 1995). This family is defined by their stable association with heterogeneous nuclear RNA, hence the name hnRNPs. These proteins are found in complexes that form major nuclear structures and are as abundant as histones. More than 50 hnRNPs have been identified in human cells. However, the function of these proteins and the ribonucleoprotein structures they form in the nucleus are not clearly understood.

The most well characterized of these proteins are the human hnRNPs, of which there are ~20 major hnRNP proteins, referred to as hnRNP A1 through U. Again, they are predominantly nuclear in subcellular distribution and bind to pre-mRNA and mRNA *in vivo*. They contain characteristic RNA-binding motifs, which can include various combinations of the following: 1) ribonucleoprotein consensus sequence RNA-binding domain (RBD), RNA recognition motif (RRM) or RNP motif, 2) arginine-glycine-glycine (RGG) box motif, and 3) hnRNP K-homology (KH) domain.

Metazoan hnRNPs have been classified by their RNA-binding motifs and subclasses often exhibit similar RNA-binding preferences (reviewed in Dreyfuss et al., 1993, and Swanson, 1995). In addition, subclasses of metazoan hnRNPs can consist of multiple immunologically related isoforms, at least some of which are generated by alternative splicing. The most abundant subclass of metazoan hnRNPs is the A/B-type proteins. There are multiple isoforms of these proteins in human cells, and in fact, this subclass accounts for all of the structurally characterized hnRNPs in *Drosophila*. The hnRNP A/B proteins share common structural features including two amino-terminal RBD motifs and a glycine-rich carboxyl terminus, which often contains an RGG box

motif. Similarly, the hnRNP D proteins also consist of multiple isoforms that possess these general structural characteristics. In contrast, the hnRNP C proteins contain a single RBD, while other hnRNPs contain as many as three or four RBDs. Only a few of the major human hnRNP proteins do not contain an RBD. These include hnRNP K and J, which contain the KH motif, and hnRNP U, whose RNA-binding activity resides in an RGG box motif.

In addition to these RNA-binding motifs, hnRNPs contain so-called auxiliary domains. It is postulated that these auxiliary domains are responsible for protein-protein interactions. The existence of a variety of RNA-binding motifs and their multiplicity in many hnRNPs suggests that hnRNPs are capable of recognizing and binding to specific regions of pre-mRNAs and mRNAs. This binding capacity, in combination with potential protein-protein interactions between hnRNPs, may facilitate and maintain packaging of pre-mRNAs and mRNAs and also allow distal RNA sequences to be brought into proximity within an hnRNA-hnRNP complex during pre-mRNA processing and mRNA transport.

Functions of Pre-mRNA/mRNA-Binding Proteins

Nuclear Functions

Single-stranded DNA/RNA binding and annealing

The first functional activity described for hnRNPs was the ability to bind to single-stranded DNA and RNA and destabilize helices, originally discovered by the isolation of UP1, a proteolytic fragment of hnRNP A1 consisting primarily of the amino-terminal RBDs (Chase and Williams, 1986). However, the full-length A1 protein does

not possess the same ability to lower the T_m of double-stranded DNA or RNA, but rather exhibits strand annealing activity between complementary sequences, an activity which lies within the carboxyl terminus of A1 (Pontius and Berg, 1990; Munroe and Dong, 1992). This is a common feature of hnRNPs and RNA annealing activity from HeLa cells co-purifies with about half of the major HeLa hnRNPs (Portman and Dreyfuss, 1994). The predominance of this conserved activity among hnRNPs suggests that it is fundamental to a general function shared by many hnRNP proteins. Such an activity may afford hnRNPs the ability mediate RNA•RNA base pairing to facilitate or prevent the formation of RNA structures in a manner conducive to pre-mRNA processing.

Nascent transcript packaging

Immediately following the initiation of transcription by pol II, hnRNPs associate with pre-mRNAs to form large macromolecular structures (reviewed in Dreyfuss et al., 1993, and Swanson, 1995). One of the earliest models proposed for the function of hnRNPs was the ribonucleosome model. According to this model, hnRNPs associate with and package nascent pol II transcripts in a manner analogous to the packaging of DNA by histones into nucleosomes, thereby serving a general function in managing and protecting pol II transcripts. When the soluble nuclear fraction of vertebrate nuclei, or nucleoplasm, is isolated under conditions that minimize RNase activity, hnRNPs sediment with hnRNA as complexes with a sedimentation coefficient >200S. These large complexes consist of many RNA-binding proteins, including the 20 major hnRNPs. In contrast, when nucleoplasm is mildly treated with RNase, 30-40S particles are isolated, composed primarily of six 'core' hnRNPs. These observations suggested that these hnRNPs bind to hnRNA, independently of the RNA sequence, to form macromolecular

complexes that protect the hnRNA from RNase degradation during processing. However, the existence of these complexes as functional units *in vivo* has not been established.

A second model for hnRNP function is that hnRNP proteins associate with nascent pre-mRNAs to form transcript-specific complexes. Perhaps the strongest evidence for this model comes from the visualization of transcripts from individual genes in *Drosophila* salivary gland chromosome spreads (Amero et al., 1992; Matunis et al., 1993). By using antibodies specific for a variety of hnRNPs, it has been demonstrated that the same set of hnRNPs can bind to a number of different transcripts. However, the stoichiometry and overall pattern of hnRNP binding varies from one transcript to another. This model is further supported by a number of studies using different techniques that demonstrate clear sequence-preference binding by a variety of hnRNPs (reviewed in Dreyfuss et al., 1993). Still, the formation of transcript-specific complexes is probably not due solely to sequence-preference binding by hnRNPs. Because of the abundance of hnRNPs, the number of molecules of any given hnRNP protein probably far exceeds the number of discrete sites available for sequence-preference binding. As general single-stranded nucleic acid binding proteins, hnRNPs may occupy lower affinity sites as well. Subsequent protein-protein interactions between hnRNPs may lead to higher order structures and confer overall stability to complexes. Therefore, the packaging of pol II transcripts may combine features of both the ribonucleosome and transcript-specific models, resulting from the sequence of the hnRNA, the binding preferences and relative abundance of hnRNP proteins, competition between proteins for binding sites, and interactions among the hnRNPs bound to a single transcript.

The potential for the formation of transcript-specific hnRNP complexes also suggests that this RNP structure may serve as the substrate for transcript-specific pre-mRNA processing, which must occur in a precise manner particular to any given pre-mRNA. This has been referred to as the substrate presentation model and is supported by structural studies. The interaction between specific hnRNP RBD motifs with a preferred RNA substrate has revealed that the RBD appears to serve as a platform to which the RNA binds and remains exposed rather than buried within the RBD (Görlach et al., 1992; Jessen et al., 1992). This would allow the RNA to remain accessible to pre-mRNA processing factors, suggesting a substrate-presentation role for hnRNPs. Such substrate presentation may affect pre-mRNA processing directly or indirectly.

Pre-mRNA splicing

One of the first proposed roles for hnRNPs in pre-mRNA processing was in splicing. *In vitro*, splicing could be inhibited by the addition of antibodies against the hnRNP A/B and C proteins (Sierakowska et al., 1986). Also, immunodepletion of the C proteins blocked splicing (Choi et al., 1986). However, it has not been possible to restore splicing activity by the addition of the respective recombinant proteins following immunoinhibition or immunodepletion. Therefore, it is possible that removal of these abundant proteins from splicing extracts also results in the loss of specific essential splicing factors, leading to the observed splicing defect. More recently, selection-amplification studies starting with random RNA sequences have suggested a role for hnRNP A1 in splicing, with the majority of selected sequences resembling 3' splice sites (Burd and Dreyfuss, 1994). Also, hnRNP I has also been identified as the polypyrimidine tract binding protein (Gil et al., 1991; Patton et al., 1991; Ghetti et al.,

1992), a protein which can be specifically crosslinked to the polypyrimidine tract. Mutations in the polypyrimidine tract that reduce the crosslinking of hnRNP I/PTB also decrease the efficiency of spliceosome formation (Gil et al., 1991). However, neither hnRNP A1 nor hnRNP I has been shown to be an essential splicing factor.

In contrast, there is substantial evidence for the involvement of hnRNPs in splice site selection during alternative splicing. The choice of 5' splice sites is clearly influenced by the ratio between two RNA-binding proteins, hnRNP A1 and the SR protein ASF/SF2 (alternative splicing factor/splicing factor 2) (Mayeda and Krainer, 1992). This has been observed both *in vitro* and *in vivo* (Cáceres et al., 1994), with higher concentrations of A1 favoring distal 5' splice sites and higher concentrations of ASF/SF2 favoring proximal 5' splice sites.

Formation of 3'ends

Specific RNA sequence elements are also required for 3'-end processing of pre-mRNAs. The existence of such sequences suggested that there may be pre-mRNA binding proteins that bind to these elements. Several RBD-containing proteins are required for mammalian 3'-end formation, including PAP (Martin and Keller, 1996), PAB II (Nemeth et al., 1995), the 64 kDa subunit of CstF (Takagaki et al., 1992), and the 68 kDa subunit of CF I_m (Rüegsegger et al., 1998). With the exception of PAP, these proteins bind or crosslink specifically to RNA substrates containing 3'-end processing sequences. Likewise, a few components of the essential factors for yeast 3'-end formation are RBD-containing proteins, including Pab1p (Sachs et al., 1986), Rna15p (Minvielle-Sebastia et al., 1991) and Hrp1p (Kessler et al., 1997). Based on its ability to bind to cytoplasmic polyadenylate tails, Pab1p may have the capacity to bind to the

elongating tail during poly(A) addition, analogous to PAB II. While both Rna15p and Hrp1p can be crosslinked to RNA substrates for *in vitro* 3'-end formation, the specificity of these interactions has not been clearly defined. Still, none of these RBD-containing proteins had been characterized as hnRNPs.

The hnRNP C proteins bind polyuridine stretches and photocrosslink to uridylate tracts near the 3'-end cleavage site (Wilusz and Shenk, 1990), but a direct role for a hnRNP protein in 3'-end processing has not yet been demonstrated. However, alternative usage of polyadenylation sites is an important regulatory mechanism during gene expression. Therefore, it is possible that hnRNPs may function in cleavage site selection in a manner analogous to the regulation of splice site selection by hnRNP A1 and ASF/SF2.

Messenger RNA export

As described previously, the substrates for mRNA export are RNP complexes, thereby implicating the associated hnRNPs as potential mediators and regulators of RNA export. Once thought to be exclusively nuclear proteins, it has been observed that many hnRNP proteins continuously shuttle between the nucleus and the cytoplasm of cells and in fact, these shuttling hnRNPs can be found associated with cytoplasmic mRNA (Piñol-Roma and Dreyfuss, 1992). These observations strongly supported a role for hnRNPs in the export of mRNA. More recently, the identification of a nuclear export signal (NES) within a number of these proteins, including hnRNP A1/A2 and hnRNP K, further suggested that these proteins are directly involved in mRNA export (Michael et al., 1995, 1997). These sequences can mediate rapid, temperature-dependent export of a nuclear-restricted heterologous protein, and NES-mediated export is saturable (Görlich and

Mattaj, 1996; Nakielny and Dreyfuss, 1997). Various types of nuclear export signals have been identified, suggesting the possibility of multiple pathways for RNA export, but it is likely that these pathways share at least some common components. The nuclear export signals of hnRNP A1/A2 and hnRNP K, termed M9 and KNS, respectively, are unique among the described nuclear export signals because they mediate the transport of these proteins into and out of the nucleus, serving also as nuclear localization signals (NLS). The M9 sequence interacts with transportin, a protein similar to the nuclear import factor importin β, and mediates nuclear import of M9-containing proteins (Pollard et al., 1996; Nakielny et al., 1996). By analogy, M9, or other hnRNP NESs, may be bound by similar transport factors in the nucleus to facilitate mRNA export (Izaurralde et al., 1997).

Intriguingly, there are also hnRNPs, such as the hnRNP C proteins, which do not shuttle and are actively restricted to the nucleus by a nuclear retention signal (NRS) (Nakielny and Dreyfuss, 1996). This sequence apparently overrides an NES, as fusing the hnRNP C NRS to shuttling hnRNPs blocks the export of these proteins. While the mechanism of nuclear retention is not clearly understood, it may involve high-affinity RNA-binding sites for the C proteins such as those required for pre-mRNA splicing and 3'-end formation, as described previously. Thus, it appears that mRNA export might be regulated by the dynamic association of NES- and NRS-containing hnRNPs. NRS-bearing hnRNPs may prevent the premature export of partially processed or unprocessed pre-mRNAs. Once processing is complete, NRS-containing hnRNPs would dissociate from the mature mRNA, thus allowing the NES-containing hnRNPs to function to mediate mRNA export.

Cytoplasmic Functions

With the discovery that hnRNPs are not exclusively nuclear proteins, it has been suggested that at least some hnRNPs may have cytoplasmic functions, perhaps in mRNA trafficking, translation, or stability. Many cytoplasmic mRNA binding proteins that directly affect these aspects of mRNA biogenesis have been identified in various organisms (St. Johnston, 1995; Curtis et al., 1995; Decker and Parker, 1995). However, it remains to be seen whether the same will be true for hnRNPs.

The functional link identified between transcription and translational masking suggests the possibility that hnRNPs, which associate with nascent pre-mRNAs during transcription, may be important in the assembly of translationally silent mRNPs. The hnRNP A/B-type hrp36 of *C. tentans* is an abundant protein in BR RNA-containing polysomes and a good candidate for an hnRNP with a cytoplasmic function. However, hrp36 remains associated with these polysomes during translation, which would argue against a role in translational masking.

The AUF1/hnRNP D proteins bind to A+U-rich elements (ARE) found in the 3' UTR of some proto-oncogene and lymphokine mRNAs (DeMaria and Brewer, 1996; Kiledjian et al., 1997). These elements, in part, regulate the degradation of these mRNAs, suggesting that AUF1/hnRNP D proteins may contribute to the regulation of mRNA turnover. However, a direct role for these proteins in the regulation of mRNA stability has not been established.

Identification and Characterization of hnRNPs in the Yeast *Saccharomyces cerevisiae*

The yeast *Saccharomyces cerevisiae* provides a powerful experimental system for investigating the regulation of cellular metabolism because of the ability to complement biochemical and cell biological studies with genetic analyses. This organism is especially attractive for examining the regulation of gene expression, and specifically the biogenesis of pre-mRNAs and mRNAs, because of the fundamental conservation of these pathways in all eukaryotes. In addition, the regulation of these pathways in budding yeast may be simplified in some respects, particularly in the case of pre-mRNA splicing, since the average transcriptional unit in *S. cerevisiae* is considerably shorter and less complex than those found in higher eukaryotes. The absence of alternative splicing and developmental or tissue-specific pathways in yeast should provide a more basic context in which to evaluate the essential functions of hnRNP during the regulation of gene expression.

Recently, proteins analogous to metazoan hnRNPs have been described in *S. cerevisiae*, the nuclear polyadenylated RNA binding, or Nab, proteins (Anderson et al., 1993a, 1993b; Wilson et al., 1994). These proteins are highly abundant and predominantly nuclear in distribution. They associate with poly(A)$^+$ RNA *in vivo* and contain structural motifs characteristic of metazoan hnRNPs including RNA-binding domains and auxiliary domains that are putative protein-protein interaction motifs. The original isolation of Nab proteins involved a technique that exploits the association of these proteins with poly(A)$^+$ RNA *in vivo*. Briefly, living yeast cells were UV-irradiated to induce the formation of covalent crosslinks between these proteins and the RNAs to which they are intimately associated *in vivo*. The population of proteins bound to

polyadenylated RNAs were purified by oligo(dT)-cellulose chromatography and used for the production of antibodies. These antibodies were then used to isolate the corresponding genes. This procedure resulted in the identification of three proteins that satisfied the criteria established to authenticate hnRNPs, Nab1p (also known as Npl3p/Nop3p/Mts1p/Mtr13p), Nab2p, and Nab3p.

Like their metazoan counterparts, the yeast hnRNPs are modular proteins, consisting of varying patterns of both RNA-binding motifs and putative protein-protein interaction domains. The Nab1/Npl3 protein contains a central conserved RBD/RRM as well as a more divergent RRM-like domain. The carboxyl terminus has several RGG box motifs as well as a number of SR dipeptides. In contrast, Nab2p contains a central RGG box and a unique RNA-binding domain consisting of repeats of a C_3H zinc finger-like cysteine/histidine-rich motif. Only one RNA-binding motif, a central RBD, was found in Nab3p, while the amino terminus was extremely acidic and the carboxyl terminus was composed of variable repeats of adjacent proline (P) and glutamine (Q) residues. Nab1p/Npl3p and Nab2p also possessed Q/P-rich regions, located in the amino terminus of each protein. While all of these proteins contained previously characterized RNA-binding motifs (or experimentally confirmed RNA-binding domains in the case of Nab2p), none of the Nab proteins strongly resembled metazoan hnRNPs in its overall structure. Especially curious was the lack of a yeast hnRNP bearing strong resemblance to the most abundant metazoan hnRNPs, the A/B-type proteins. More recently, it has been suggested that Nab1p/Npl3p might be structurally related to the A/B-type proteins, if the central RRM-like motif was considered to be an RBD (Henry et al., 1996).

However, there is very little sequence similarity between this family of metazoan hnRNPs and Nab1p/Npl3p.

The similarity that exists between isoforms of the various mammalian hnRNPs suggests that these proteins may have redundant functions, with each isoform able to compensate for another. Initial genetic studies revealed a very important feature of the yeast hnRNPs, namely that each of them is essential for cell viability. While they may be somewhat redundant, perhaps in their capacity for sequence-independent RNA binding, it was clear that the Nab proteins have crucial and distinct functions in cellular metabolism. While our understanding of the function of Nab proteins is far from complete, this aspect of the role of Nab proteins during cell growth has been reflected thus far in functional studies. For instance, a variety of functions have been suggested for proteins identical to Nab1p. Npl3p/Nop3p was originally implicated in nuclear protein import (Bossie et al., 1992) and pre-rRNA processing (Russell and Tollervey, 1992). More recently, it was reported that Npl3p shuttles between the nucleus and cytoplasm (Flach et al., 1994) and a role for Npl3p in mRNA export has been proposed (Flach et al., 1994; Russell and Tollervey, 1995; Singleton et al., 1995; Lee et al., 1996). Likewise, Nab2p has been implicated in mRNA export as well as the regulation of polyadenylation (Anderson, 1995). The function of Nab3p is perhaps the least understood of all the Nab proteins (Wilson, 1996). The RBD of Nab3p is most similar to the hnRNP C proteins, suggesting a possible role in pre-mRNA splicing (Wilson et al., 1994). In support of this, depletion of Nab3p from cells correlates with the accumulation of *ACT1* pre-mRNA and an increased ratio between *CYH2* pre-mRNA and mRNA. However, this defect is mild

compared to those observed due to the loss of known splicing factors, and mutations in Nab3p could not confirm that this was a direct effect.

Previous investigation of the Nab proteins has told us some important things about hnRNP proteins. First, hnRNPs have been conserved throughout evolution, and therefore are certain to play fundamental roles in maintaining cell growth and metabolism. Second, while the existence of hnRNPs in *S. cerevisiae* has been known for only a fraction of the time since hnRNPs were first discovered, we have already gained significant insights regarding their contributions to pre-mRNA processing and mRNA transport. Finally, with similarities in pre-mRNA processing and mRNA export between yeast and metazoan cells continually emerging, discovering how hnRNPs regulate gene expression will certainly be expedited by ongoing studies in this model genetic organism.

MATERIALS AND METHODS

Growth Conditions and Media

All bacterial plasmids and yeast shuttle vectors were propagated in either *E. coli* strain DH5α [*supE*44 Δ*lac* U169 (Ø80 *lacZ*ΔM15) *hsdR*17 *recA*1 *endA*1 *gyrA* 96I *thi*-1 *relA*1] or DH10B [F- *mcr*A Δ(*mrr-hsd*RMS-*mcr*BC) *lacZ*ΔM15 *lac*ΔX74 *deo*R *recA*1 *ara*Δ139 Δ(*ara,leu*) 7697 *gal*U *gal*K λ- *rps*L *end*A1 *nup*G]. Bacteria were grown in LB media (1% Bacto-tryptone, 0.5% Bacto-yeast extract, 1% NaCl, pH 7.5, 2% Bacto-agar for plates only) supplemented with 100 μg/ml ampicillin at 37°C. Yeast cells were cultured in YPD (1% Bacto-yeast extract, 2% Bacto-peptone, 2% dextrose, 2% Bacto-agar for plates only), synthetic dextrose (SD), synthetic raffinose (SR) or synthetic galactose (SG) media (0.67% Bacto-yeast nitrogen base w/o amino acids, 2% dextrose, 2% raffinose or 2% galactose, 2% Bacto-agar for plates only). Synthetic minimal media were supplemented with L-amino acids, adenine and uracil as needed (Rose et al., 1990). For mitochondrial induction, yeast cells were cultured in semi-synthetic lactate (SSL) media (0.3% Bacto-yeast extract, 0.1% potassium phosphate, monobasic, [pH 5.5, using sodium hydroxide], 0.1% ammonium chloride, 0.05% calcium chloride, dihydrate, 0.05% sodium chloride, 0.06% magnesium chloride, hexahydrate, 2% lactate) according to previously described procedures (Yaffe, 1991.) Prior to sporulation, diploid strains were grown on YPD or minimal media overnight and transferred to sporulation plates (1%

potassium acetate, 0.025 mg/ml zinc acetate) supplemented as described above. Growth of yeast strains was at 30°C on a roller drum, in shaking water baths, or in an air rotary shaker unless otherwise stated. Plasmid shuffling employing 5-fluoro-orotic acid (5-FOA) to evict plasmids carrying *URA3* was performed as described previously (Sikorski and Boeke, 1991). All of the yeast strains used in this study and their genotypes are described in the appendix.

Cell Transformations

DH5α was made electrocompetent for transformation as described previously (Sambrook et al., 1989). *E. coli* transformations were performed using a gene pulser (BioRad, Richmond, CA) according to the manufacturer's instructions. Yeast transformations were performed as described previously (Ito et al., 1983) with a few modifications. Cells were grown to an OD_{600}=0.4-1.0 in YPD or SD. For each transformation, 10 ml of cells were harvested by centrifugation for 5 min at 1,500 x *g*. The pelleted cells were washed once in TE [10 mM Tris-HCl, pH 7.5, 1 mM ethylenediaminetetraacetic acid (EDTA)] containing 0.1 M lithium acetate. Cells were centrifuged as described above and resuspended in 100 µl of TE/0.1 M lithium acetate. The DNA to be transformed (2-25 µg), carrier DNA (15 µg calf thymus DNA), and 600 µl TE/0.1 M lithium acetate/50% 3350 MW polyethylene glycol (PEG_{3350}) were added to the cells, mixed, and incubated at room temperature (RT) on a Labquake shaker for 30 min. Cells were heat-shocked at 42°C for 15 min. After cooling to RT, 400 µl of 1 M sorbitol was added, and the cells were harvested by centrifugation at 3,000 x *g* for 1 min in an Eppendorf microcentrifuge. The cell pellet was washed with 1 ml of 1 M sorbitol

and centrifuged as before. Cells were resuspended in 1 ml of 1 M sorbitol and plated on SD plates appropriate for the selection of cells containing the transformed DNA.

Yeast Genetic Manipulations

All genetic manipulations were performed as previously described (Guthrie and Fink, 1991) with the following exceptions. Matings were performed by patching haploid strains together on YPD, incubating at 24°C or 30°C for 3-5 h, and then plating cells on media for selection of diploids, or on YPD for physical manipulation of zygotes. For segregation analysis, asci were suspended in 1 M sorbitol containing 5 mg/ml Zymolyase 20T (Seikagaku Corporation, Tokyo, Japan). Asci were digested for 1-3 h at RT and streaked onto YPD plates. Dissection of tetrads into individual spores was performed by micromanipulation using a dissection microscope followed by incubation at the appropriate temperature for germination. Cells were replica plated onto selective SD plates to test for auxotrophic markers and/or YPD plates at various temperatures for determination of conditional lethal phenotypes.

Nucleic Acid Isolation Procedures

Yeast genomic DNA was isolated from cells grown to saturation in YPD. Five ml of cells were pelleted by centrifugation for 5 min at 1,250 x *g*. Cells were resuspended in 0.5 ml of 1 M sorbitol plus 100 µl 0.5 M EDTA, pH 8.0, 18 µl 1 M dithiothreitol (DTT), and 50 µl of 5 mg/ml Zymolyase 100T (Seikagaku Corp.) and incubated for 1 h at 37°C. Spheroplasted cells were harvested by centrifugation at 3,000 x *g* for 1 min in a microcentrifuge. Spheroplasts were lysed by resuspension in 500 µl 50 mM Tris-HCl,

pH 7.5, 20 mM EDTA plus 10% sodium dodecyl sulfate (SDS) to 1% final concentration and incubation at 65°C for 30 min, with intermittent mixing. Next, 200 µl of 5 M potassium acetate was added followed by mixing and incubation on ice for 20 min. The cell lysate was centrifuged at 16,000 x *g* in a microcentrifuge for 15 min at 4°C. The supernatant was transferred to a new tube and the DNA was precipitated by adding one volume of isopropanol and incubating at RT for 10 min. DNA was pelleted at 16,000 x *g* for 1 min at RT, washed with 70% ethanol and dried briefly in a Speed-Vac (Savant, Marietta, OH). The DNA was resuspended in 250 µl TE and treated with 5 µl of 10 mg/ml RNaseA at 37°C for 1 h. The lysate was extracted twice with phenol and precipitated with 1/10 volume of 4 M ammonium acetate and 2.5 volumes of 100% ethanol. After incubation on a dry ice/ethanol bath for 10 min, the DNA was recovered by centrifugation, washed and dried as described above, and resuspended in 100 µl TE.

Plasmid rescues from yeast were accomplished as described previously (Strathern and Higgins, 1991) and DNA was purified from crude lysates using a Gene Clean kit (Bio101, Vista, CA).

Total RNA was isolated from yeast cells by extraction with hot acid phenol as described previously (Guthrie and Fink, 1991). Prior to extraction, yeast cells were harvested by pouring cells over ice and pelleting by centrifugation at 2,000 x *g* for 5 min at 4°C. Final RNA pellets were resuspended in diethyl pyrocarbonate (DEPC)-treated d_2H_2O. Poly(A)$^+$ RNA was purified from 2 liters of yeast cells grown to mid-log phase (OD_{600}=1.0-3.0) unless otherwise stated. Cells were harvested by centrifugation at 2,000 x *g* at 4°C and washed once in 200-250 ml phosphate-buffered saline [PBS (0.14 M NaCl, 2.7 mM KCl, 10 mM $Na_2H_3PO_4$, 1.7 mM KH_2PO_4, pH 7.4)]. The cell pellet was

transferred to a chilled mortar to be freeze-dried and pulverized in liquid nitrogen (N_2). The frozen cells were ground into a fine powder, while continuing to add liquid N_2 as needed to keep frozen. The frozen cell powder was transferred to a Dounce homogenizer at RT containing lysis buffer [RSB100 (10 mM Tris-HCl, pH 7.4, 100 mM NaCl, 2.5 mM $MgCl_2$) containing 1% SDS, 0.5% 2-mercaptoethanol (βME), 10 mM vanadyl ribonucleoside complex, and 350 μg/ml proteinase K (Boehringer Mannheim, Indianapolis, IN)]. After 12 strokes with pestle A, the cell lysate was split to two 50 ml conical tubes and incubated at 42°C for 30 min in a circulating water bath. After the addition of 0.5 M EDTA to a final concentration of 10 mM, incubation was continued at 42°C for 15 min. Lysates were then transferred to 65°C for 10 min, chilled rapidly in an ice water bath, and 10 M lithium chloride added to 0.5 M final concentration. Cell debris was removed by centrifugation at 4,500 x *g* for 5 min at 4°C. The supernatant was batch-bound to 0.2 g of oligo(dT)-cellulose (Gibco/BRL) in binding buffer (10 mM Tris-HCl, pH 7.4, 1 mM EDTA, 0.5% SDS, 0.5 M LiCl) for 30 min at RT with gentle mixing. The oligo(dT)-bound poly(A)$^+$ RNA was collected by column chromatography and washed thoroughly with binding buffer. The poly(A)$^+$ RNA was eluted from the column in 10 ml of elution buffer (10 mM Tris-HCl, pH 7.4, 1 mM EDTA, 0.05% SDS), heated to 65°C for 10 min, and chilled rapidly as above. Poly(A)$^+$ RNA was re-selected on the oligo(dT) column, washed, and eluted in 1 ml fractions. Peak poly(A)$^+$ RNA-containing fractions, as determined spectrophotometrically by A_{260}, were pooled, extracted once with phenol and precipitated with 1/10 volume 3 M sodium acetate and 2 volumes of 100% ethanol. Prior to use, RNAs were pelleted, resuspended in DEPC-treated d_2H_2O to be phenol extracted again, and re-precipitated as above.

Small-scale plasmid preparations were performed by alkaline lysis (Sambrook et al., 1989). Large-scale preparations of plasmid DNA were performed using the Wizard Midi-prep plasmid kit (Promega, Madison, WI). Recombinant DNA plasmids were constructed using conventional cloning techniques (Sambrook et al., 1989) and descriptions for all of the plasmids used in this study can be found in the appendix.

Isolation of UV Cross-linked Polyadenylated RNA-RNP Complexes and Antibody Preparation

UV cross-linking and isolation of poly(A)$^+$ RNA-protein complexes was performed according to a previously described procedure (Anderson et al., 1993b) using a total of 20 L of cells at OD_{600}=2-3. Polyclonal antisera against polyadenylated RNPs were raised by immunizing BALB/c mice with UV cross-linked RNPs isolated from YJA213 grown at 30°C, as described previously (Anderson et al., 1993b). Polyclonal antisera were also prepared against a fusion protein containing the glutathione-S-transferase (GST) protein and a partial *NAB4* open reading frame (nt 97 to 1602). The GST-Nab4 fusion protein was produced in JM105 cells from an expression plasmid, pRNPex4.7, constructed using pGEX-4T-1 (Pharmacia, Piscataway, NJ), and purified according to the manufacturer's protocols. Mice were injected three times with 0.2 μg of antigen per injection. All antibodies generated for this study were prepared by the University of Florida Interdisciplinary Center for Biotechnology Research (ICBR) Hybridoma Laboratory.

NAB4 Gene Isolation and Sequence Analysis

The *NAB4* gene was isolated by immunoscreening a λgt11 genomic expression library as previously described (Snyder et al., 1987), using a 1:400 dilution of the polyclonal antisera elicited against the cross-linked RNPs described above. The full length *NAB4* gene (pNAB4.0) was isolated by hybridization screening of a YCp50 genomic library (Yeast Genetic Stock Center, Berkeley, CA) using an 800 bp EcoRI/HincII fragment from pRNP4.0. Colony filters were hybridized in 50% formamide/5 SSC (1X SSC=0.15 M NaCl/0.015 M sodium citrate, pH 7.0)/0.2% SDS/5X Denhardt's soultion/100 μg/ml salmon sperm DNA at 42°C. Filters were washed in 0.5X SSC/0.1% SDS at 65°C for 30 min.

All DNA sequences were determined for both strands using restriction fragments cloned into pSP72 (Promega) and SP6/T7 or gene-specific oligonucleotide primers as described (Wilson et al., 1994). The sequences of the oligonucleotides used in this study, for sequencing and DNA amplification by polymerase chain reaction are listed in the appendix. DNA sequencing was performed manually and the University of Wisconsin Genetics Computer Group programs were used to analyze all sequence information. Database searches were performed using the BLAST network service at the National Center for Biotechnology Information (Altschul et al., 1992). The GenBank accession number for *NAB4* is U35737.

DNA and RNA Blot Analysis

Yeast genomic DNA (10-20μg) was treated with restriction endonucleases, separated by agarose gel electrophoresis and transferred to Hybond-N^+ nylon membrane

(Amersham Corporation., Arlington Heights, IL) by capillary blotting (Sambrook et al., 1989). Total RNA samples were treated with dimethylsulfoxide (DMSO)/glyoxal, fractionated on 1.2-1.4% agarose gels (Sambrook et al., 1989), and transferred to Hybond-N^+ in 20X SSC. Hybridization was performed in a hybridization oven (Robbins) at 65°C overnight (12-16 h) in a solution containing 1% bovine serum albumin (BSA), 1 mM EDTA, 25 mM Na_2HPO_4, and 7% SDS. After hybridization, blots were washed once each in 2X SSC/0.1% SDS followed by 1X SSC/0.1% SDS at RT for 15 min followed by one wash each in 1X SSC/0.1% SDS and 0.5X SSC/0.1% SDS at 65°C for 15 min. Radiolabeled probes were prepared using a Random Primed Labeling Kit (Gibco/BRL) and [α^{32}P]-dCTP (3000Ci/mmol, Amersham Corp.). Probes used for RNA blot analysis were: (a) *ACT1*, representing the intron (nt 21 to 279, where ATG start codon is nt 1 to 3) or the intron and second exon (nt 21 to 724); (b) *CYH2*, representing the full-length gene (nt –272 to 1065); (c) *CUP1*, representing the full-length gene (-285 to 477); and (d) *TPI1*, a HinDIII-EcoRV restriction fragment (nt 241 to 726).

Yeast Total Cell Protein Preparation

Yeast cells were grown to an OD_{600}=1 and 5 ml of cells were chilled on ice and harvested by centrifugation at 2,000 x *g* for 5 min at 4°C. The cell pellet was resuspended in 0.2 ml of ice-cold 10% trichloroacetic acid (TCA), to which acid-washed glass beads were added to just below the meniscus. Cells were lysed by vortexing five times for 15 s with cooling on ice in between vortexing. The liquid was aspirated and transferred to a new tube and the beads were washed with 200 µl of 10% TCA. This liquid was removed and pooled with the first aspirate. Precipitable proteins were pelleted

in a microcentrifuge at 16,000 x *g* for 10 min at 4°C. The supernatant was removed and the pellet was resuspended in Laemmli buffer, neutralized with a minimal volume of 1 M Tris-base, and boiled for 3 min prior to fractionation by SDS-polyacrylamide gel electrophoresis.

RNA Binding Studies, Gel Electrophoresis and Immunoblotting

In vitro transcription/translation of pSP72-based expression plasmids were performed using T7 RNA polymerase and rabbit reticulocyte lysates according to the manufacturer's instructions (Promega). *In vitro* binding reactions of ^{35}S-methionine labeled Nab4p or hnRNP A2 protein to RNA homopolymers were performed as described (Anderson et al., 1993b). Proteins were resolved by SDS-polyacrylamide gel electrophoresis (SDS-PAGE) using 12.5% (final acrylamide concentration) separation gels. Coomassie Blue-stained gels were impregnated with 2,5-diphenyloxazole for visualization of ^{35}S-labeled proteins by fluorography. For immunoblotting, proteins were transferred to nitrocellulose following electrophoresis using a semi-dry electroblotting apparatus (Bio-Rad Laboratories, Hercules, CA) for 0.5-1 h at 15-20V. Immunoblotting was performed as described (Anderson et al., 1993b; Wilson et al., 1994) using the following antibody dilutions: 3H1 (Nab4p), 1:500; 2B1 (Pub2p), 1:10,000; 3F2 (Nab2p), 1:500; 1G1 (Pab1p), 1:1000; 1E4 (Nab1p), 1:500; anti-Npl3p rabbit polyclonal antisera (a gift from C. Siebel and C. Guthrie, University of California, San Francisco, CA), 1:1000. Reactive antigens were visualized by enhanced chemiluminescence using horseradish peroxidase-conjugated sheep anti-mouse or donkey anti-rabbit secondary antibodies at 1:5,000 (Amersham Corp.) with film exposures of 15 s-1 min.

Indirect Cellular Immunofluorescence

To determine the subcellular localization of cellular proteins, the following procedure was used. Yeast cells were grown to OD_{600}=0.1-0.5 in the desired media. For each strain or condition examined, $1.5X10^8$ cells were fixed by adding 10% formaldehyde [freshly prepared from paraformaldehyde (EM grade, Polysciences, Warrington, PA) in 100 mM KH_2PO_4, pH 6.5] to a final concentration of 4%. Following incubation for 2 h at RT with gentle mixing, cells were pelleted at 800 x *g* for 2.5 min at RT. Pellets were washed twice in 25 ml of WB1 (100 mM KH_2PO_4, pH 6.5) and once with 25 ml of WB2 (100 mM KH_2PO_4, pH 6.5/1.2 M sorbitol), pelleting cells by centrifugation in between washes as described above. Cells were resuspended in 1 ml of SB (100 mM KH_2PO_4, pH 6.5/1.2 M sorbitol/30 mM βME) and 30 µl of Zymolyase 100T (Seikagaku), 10 mg/ml in WB2, was added. Cells were incubated for 30-45 min until >90% spheroplast formation with the degree of cell wall digestion monitored by phase-contrast microscopy (400X magnification). Spheroplasts were pelleted at 3,000 x *g* for 1 min at RT in a microcentrifuge and washed gently in 1 ml of WB2. Cells were resuspended in 1 ml of WB2 and cells were adhere to polylysine-coated 10-well HTC Blue slides (Cel-Line Associates, Newfield, NJ) for 10 min at 4°C, 10 µl/well. Slides were washed once in ice-cold PBS and sequentially incubated at -20°C in 100% methanol for 5 min and 100% acetone for 30 s. Slides were then washed three times with PBS, incubated in 0.1% Triton X-100 in PBS, and washed again, three times in PBS. Prior to antibody binding, each slide well was blocked with 10 µl of 3% BSA in PBS for 30 min at RT. Monoclonal antibodies were diluted in 3% BSA in PBS as follows: 3F2, 1:500;

1G1, 1:5,000; 3H1, 1:500; 2D1, 1:500. Cell wells were incubated with 10 µl of primary antibody for 60 min at RT. Primary antibodies were detected using fluorescein- or rhodamine-conjugated goat anti-mouse IgG1 or IgG2a subclass-specific antibodies (Cappel) diluted 1:10 in 3% BSA in PBS. Each cell well was incubated with 10µl of secondary antibody for 30 min at RT. Slides were washed three times in ice-cold PBS followed by a 10 min staining with 0.5 µg/ml 4'6-diamidino-2-phenylindole (DAPI, Sigma, St. Louis, MO), and washed three more times in PBS. Slides were dried around the wells, mounting media [1 mg/ml p-phenylenedianine (Sigma) in 90% glycerol] was applied, and slides were sealed with a coverslip using clear nail polish. Fluorescent images were produced using a Nikon Optiphot-2 microscope equipped with a 100X fluorescence/differential interference contrast (DIC) objective or by digital fluorescence microscopy as described previously (Wilson, 1996).

Construction and Subcellular Distribution of GFP-Nab4p and Deletion Analysis

The coding region of the red-shifted GFP variant GFP-S65T was amplified from the plasmid pS65T-CI (Clontech, Palo Alto, CA) using the oligonucleotides MSS5110 and MSS5111. This fragment was subcloned into the SstI site of pNAB4.46 near the 5' end of the *NAB4* coding region to give pNAB4.76. Full-length GFP-Nab4p and carboxy-terminal truncated fusion proteins were produced in yeast from plasmids further described in the appendix. GFP-Nab4p fusion proteins were visualized in living cells by fluorescence microscopy using the fluorescein channel. Alternatively, cells were first washed in PBS, incubated in 0.2ug/ml DAPI in PBS for 10 min, washed again in PBS,

resuspended in mounting media diluted 1:3 in PBS, and mounted on plain glass slides sealed with a coverslip and clear tape to be visualized by digital microscopy.

Two-Hybrid Screen and Determination of β-Galactosidase Specific Activity

To identify proteins that interact with Nab4p in *vivo*, the MATCHMAKER™ Two-Hybrid System was used according to the manufacturer's protocols (Clontech) with a few modifications. The coding region of *NAB4* was subcloned into pGBT9 as an EcoRI-BamHI fragment from pNAB4p.0. This plasmid, pGBTNAB4, was transformed into the host yeast strain Y190 (Harper et al., 1993) and production of the fusion protein in YAO415 was confirmed by immunoblot analysis using mAb 3H1. YAO415 was transformed with a yeast cDNA library, generously provided by John P. Aris (University of Florida, Gainesville, FL). Interacting proteins were identified by sequencing cDNA plasmid clones rescued from His^+ β-galactosidase$^+$ transformants and the relative strength of the protein-protein interactions was measured by determining the specific activity of β-galactosidase (β-gal) using a quantitative liquid assay (Drazinic et al., 1996).

Overexpression of Nab4p

To control the expression of Nab4p, the gene was cloned behind the inducible/repressible *GAL1* promoter (pNAB4.10). This plasmid was transformed into YAO403 (see below) and plated for selection on SD media lacking uracil. The Ura^+ YAO407 and a control cell YAO408 were maintained on SR media lacking uracil. Overproduction of Nab4p was induced by streaking cells onto SG plates lacking uracil or

by the addition of 20% galactose to cells cultured in SR media lacking uracil (2% galactose, final concentration).

In Situ Hybridization and Cellular Immunofluorescence

The subcellular distribution of poly(A)$^+$ RNAs was examined by in situ hybridization of digoxigenin-labeled dT_{50} (Amberg et al., 1992) using a modification of a protocol kindly provided by A. de Bruyn Kops and C. Guthrie (University of California, San Francisco, CA). Cells were grown to OD_{600}=0.5-1.0, and 1.5 X 10^8 cells were fixed, washes, digested, and washed again as for indirect cellular immunofluorescence, and resuspended in 0.5 ml of WB2. Cells were then allowed to adhere to polylysine-coated 10-well HTC Blue slides for 15 min at 4°C, 15 µl/well. Slides were washed once with phosphate-buffered saline (PBS), incubated sequentially at -20°C in 100% methanol for 5 min and 100% acetone for 30 s, and then dried in a Speedvac (Savant) for 1 min. Slides were pre-warmed to 37°C in a humidified chamber. The hybridization solution was prepared using DEPC-treated d_2H_2O and contained 10% dextran sulfate/5X SSC/1X Denhardt's solution/33% deionized formamide/10 mM vanadyl ribonucleoside complex/100 µg/ml denatured salmon sperm DNA/200 µg/ml tRNA/0.68 µg/ml digoxigenin-labeled dT_{50}. Digoxigenin was conjugated to dT_{50} as described previously (Amberg et al., 1992). Hybridization was for 12-14 hr at 37°C followed by washes to remove excess probe as described previously (Amberg et al., 1992). Individual slide wells were sequentially incubated with 15 µl of the following solutions for 30 min at 37°C: (1) 3% BSA in PBS (antibody-binding buffer, AB); (2) anti-digoxigenin monoclonal antibody (IgG1 sub-class, Boehringer-Mannheim) diluted 1:50 in AB; (3)

fluorescein-conjugated goat anti-mouse IgG1 subclass-specific antibody (Southern Biotechnology Associates, Birmingham, AL) diluted 1:10 in AB. Finally, DNA was stained by treating cells with DAPI as described above. Following each of these steps, slides were washed three times in 250 ml of PBS. Slides were then mounted, viewed using a Nikon Optiphot-2 microscope equipped with a 100X fluorescence/differential interference contrast objective, and photographed using TMAX-100 film (Kodak, New Haven, CT).

Chromosomal Deletion of *NAB4* and Generation of *nab4* Mutant Strains

A *NAB4* null allele was constructed by replacing ~40% of the *NAB4* coding region with *LEU2* (pNAB4.9, see Appendix). This plasmid was linearized by restriction enzyme digestion within the remaining *NAB4* sequence and transformed into yJA501 to generate a leucine prototroph, YAO403. After transformation of YAO403 with pNAB4.7 to yield YAO404, the haploid YAO404-1B was isolated by sporulation and tetrad dissection. To obtain a Trp^- haploid strain, YAO404-1B was mated to S150-2B to yield YAO405. This diploid was sporulated and dissected and the Trp^- haploid YAO405-2A was used as the host strain to isolate *nab4* mutants.

Mutant *nab4* alleles were created by random PCR mutagenesis as described previously with minor modifications (Muhlrad et al., 1992). Briefly, the entire *NAB4* coding region was amplified by PCR for 30 cycles under mutagenic conditions using primers MSS194 and MSS177, 2 mM magnesium chloride, and Taq polymerase. Four separate reactions were performed, each lowering one of the four nucleotides to 40 μM with the remaining nucleotides at 200 μM. Manganese chloride was included at 15 μM

(Mn^{2+}/Mg^{2+} ratio about 1:13). Mutagenic PCR products were pooled and amplified under normal PCR conditions using MSS192 and MSS181. This PCR product was then purified by agarose gel electrophoresis and co-transformed with gapped and gel-purified pNAB4.8 into YAO405-2A. Gapped pNAB4.8 was generated by digesting the plasmid with AatII and AgeI, eliminating ~40% of the *NAB4* coding region. Transformants were replicated onto plates containing 5-FOA to select against pNAB4.7, and incubated at 24°C or 36°C. Putative mutant alleles were recovered by plasmid rescue, sequenced, and subcloned into YCp50. These YCp50 clones were transformed into the diploid strain YAO403 (*NAB4/nab4::LEU2*) and tested for dominance or recessiveness to *NAB4* at 36°C. Diploid cells harboring each of the mutants were sporulated and dissected at 24°C. Leu^+Ura^+ haploid cells containing *nab4-1*, *nab4-4* or *nab4-7* were Ts^- at 36°C.

The *nab4-9* allele was generated by PCR amplification using mutagenic oligonucleotides. First, 5' and 3' mutagenized fragments were amplified from pNAB4.0 using MSS192 and MSS565 or MSS222 and MSS564, respectively. Next, these 5' and 3' *nab4* fragments were used as template in a secondary PCR reaction using Vent DNA polymerase. A fragment of this amplified product containing the desired nucleotide changes was subcloned into pNAB4.46 to give pNAB4.95 (see Appendix). Sporulation and dissection of the diploid strain YAO467, generated by transformation of YAO403 with pNAB4.95, yielded the *nab4-9* haploid strain YAO467-1A.

The *nab4-1* allele was integrated into the chromosomal genome of two different strains. First, *nab4-1* was subcloned into pRS304, a yeast shuttling vector that lacks sequences required for episomal maintenance in yeast, to generate pNAB4.36. The XbaI site in the polylinker was eliminated by digestion of pNAB4.36 with restriction enzymes

NotI and BamHI followed by re-ligation to generate pNAB4.42. This construct was linearized within the *TRP1* gene by restriction enzyme digest with XbaI and transformed into YAO405-2A to target integration into the *TRP1* locus by homologous recombination. Trp^+ transformants were passed on 5-FOA to evict pNAB4.7 and tested for the Ts^- phenotype at 36°C. The integration of *nab4-1* at the *TRP1* locus was confirmed by Southern analysis of genomic DNA from YAO420. Second, *nab4-1* was subcloned into the yeast integrative plasmid YIp5 to generate pNAB4.45. This construct was linearized within *nab4-1* by restriction enzyme digest with BstEII and transformed into LDY133 to target integration into the *NAB4* locus by homologous recombination. Ura^+ transformants were grown in YPD overnight to allow loss of the integrated *URA3* marker. These cultured cells were plated on 5-FOA to select for loss of *URA3* then tested for the Ts^- phenotype. To confirm the integration of *nab4-1* at the *NAB4* locus the plasmid pNAB4.7 was digested with AatII and AgeI, gel purified, and transformed into the Ts^- strain YAO421. The plasmid, repaired by homologous recombination from the chromosomal *nab4-1* allele was rescued from yeast cells and sequenced, revealing only those point mutations found in *nab4-1*. Southern analysis of genomic DNA from YAO421 confirmed the integration of a single copy of *nab4-1*.

Determination of Poly(A) Tail Lengths

The length distribution of poly(A) tails was determined by following a previously described procedure with a few modifications (Minvielle-Sebastia et al., 1991). Briefly, 1 μg of total yeast RNA was 3'-end labeled in a 30 μl reaction containing 50 mM HEPES, pH 8.3, 5 μM ATP, 10 mM $MgCl_2$, 3.3 mM DTT, 10 % DMSO, 300 μg/ml

acetylated BSA (New England Biolabs), 40 μCi of ^{32}P-pCp (Amersham Corp.) and 10U of T_4 RNA ligase (New England Biolabs. Beverly, MA) for 16-18 h on ice at 4°C. Non-poly(A)$^+$ RNA was digested by the addition of 30 μg yeast tRNA, 80U RNase T1 (Gibco/BRL), 4 μg RNaseA, and 10 mM Tris-HCl, pH 7.5/0.3 M NaCl in a final volume of 80 μl. After incubation at 37°C for 2 h, digestion was stopped by the addition of 20 μl of 2 mg/ml proteinase K in 130 mM EDTA/2.5% SDS and incubation at 37°C for 30 min. Reactions were brought up to 200 μl with DEPC-treated d_2H_2O and extracted with an equal volume of phenol. Poly(A) tails were precipitated with 5 μg of RNase-free glycogen (Boehringer Mannheim) and two volumes of 100% ethanol, pelleted for 10 min at 16,000 x *g* at 4°C in a microcentrifuge, washed quickly in -20°C 70% ethanol, dried briefly, and resuspended in DEPC-treated d_2H_2O. To eliminate electrophoresis artifacts due to the glycogen, 1/60 of the reactions plus an equal volume of loading cocktail (10 mM EDTA, pH 8, in deionized formamide, no dye) were separated on 8% denaturing polyacrylamide gels. Gels were pre-run for ~ 1 h and samples were heated at 95°C for 5 min and quenched on ice prior to loading. Gels were run at 50 watts for 1.5 h and fixed two times in fresh 5% methanol/5% acetic acid.

Primer Extension Analysis

To map the site of transcription initiation for *ACT1* mRNAs, ~200 ng of the *ACT1* reverse primer MSS513 was denatured in d_2H_2O in a total volume of 14 μl at 70°C for 1 min, and cooled on ice. This primer was end-labeled at 37°C for 30 min in a 25 μl reaction containing 1X T4 DNA polynucleotide kinase buffer (50 mM Tris-Cl, pH 7.5, 10

mM $MgCl_2$, 5 mM DTT, 2 mM spermidine, 200 μM EDTA), 5 μl of [γ-^{32}P]-ATP (3000Ci/mmol, Amersham Corp.) and 1 μl of T4 DNA polynucleotide kinase (Promega). The reaction was stopped by the addition of EDTA to 20 mM. To separate the labeled oligo from the remaining free radioisotope, 175 μl of TE was added to the stopped reaction and centrifuged through a G-25 Sephadex spin column, prepared in a 1 ml disposable syringe, at 600 x *g* for 1.5 min. Primer extension was performed on 1 μg of poly(A)$^+$ RNA denatured with ~10 ng of labeled MSS513 in a total of 24 μl at 70°C for 10 min. After cooling on ice, primer extension reactions were assembled in 40 μl reactions containing 1 X PCR buffer (Gibco/BRL), 2.5 mM $MgCl_2$, 0.5 mM dGTP, dATP, dTTP and dCTP, 10 mM DTT, 2 μl of RNase inhibitor (Gibco/BRL), and 2 μl of Superscript II reverse transcriptase (Gibco/BRL). Prior to the addition of the reverse transcriptase, reactions were pre-incubated at 42°C for 5 min. Complete reactions were then incubated at 42°C for an additional 50 min. Reactions were stopped at 70°C for 15 min and then treated with 1 μl of RNase H (Gibco/BRL) at 37°C for 20 min. Following phenol extraction and precipitation with sodium acetate and 2.5 volumes of ethanol, the DNA was resuspended in 3 μl of d_2H_2O. An equal volume of Sequenase stop buffer (Amersham Corp.) was added and reactions were separated on an 8% denaturing polyacrylamide gel at 50 watts for 2 h.

Construction of cDNA Libraries and Rapid Amplification of 3'-Ends

To map the 3' termini of mRNAs, the Marathon™ cDNA Amplification Kit was used according to the manufacturer's protocols (Clontech). For *ACT1*, cDNA libraries

were generated from 1 μg of poly(A)$^+$ RNA isolated from YAO417-1D (*NAB4*) or YAO418-1D (*nab4-1*). For the *in vitro* cleavage products of *CYC1*, the full-length ^{32}P-UTP-labeled precursor was transcribed from EcoRI-cut pG4-CYC1 and purified from a 10% polyacrylamide/7 M urea gel (Preker et al., 1995). This precursor was incubated under cleavage conditions in the presence of CF II and partially purified CF IA, which contained a low amount of Nab4p, and the products were separated and purified from a 6% polyacrylamide/8.3 M urea gel (R.E. Hector and M.S. Swanson, personal communication). The four cleavage fragments were used individually to generate cDNAs. *ACT1* and *CYC1* 3' ends were amplified using the Advantage® cDNA PCR Kit and cDNA Polymerase Mix according to the manufacturer's protocols (Clontech). Gene-specific forward primers were MSS553 for *ACT1*, MSS569 for *CYC1* CP I-CP III, and MSS574 for *CYC1* CP IV. The appropriately-sized amplified products were purified from 1.6% agarose gels, digested with BamHI and EcoRI for *ACT1* and *CYC1* CPIV or EcoRV and EcoRI for the remaining *CYC1* fragments, re-purified by agarose gel electrophoresis, and subcloned into pSP72 for sequencing.

RESULTS

Objective of These Studies and Specific Aims

The ultimate goal of my research is to elucidate the pathways of gene expression required for pre-mRNA processing and the subcellular distribution of mRNA. The working hypothesis is that the transcript-specific association of hnRNPs with pre-mRNA and mRNA ensures the fidelity of these processes. It is these pre-mRNP/mRNP complexes that serve as the substrate for the various post-transcriptional modifications to pre-mRNAs. Within these RNP complexes, hnRNPs facilitate the dynamic structure of the RNA and recruit additional trans-acting factors required for the precise execution of pre-mRNA processing and nucleocytoplasmic transport of mRNA. To investigate the functions of hnRNPs in these pathways *in vivo*, I have employed the yeast *Saccharomyces cerevisiae* as a model system.

It is necessary to identify the proteins that mediate the steps of RNA processing in order to understand how these events are executed. Logical candidates for such factors are hnRNPs. The existence of additional hnRNPs in yeast seemed likely because of the vast array of hnRNPs found in higher eukaryotic cells, the similarities between the yeast Nab proteins and metazoan hnRNPs, and the fundamentally conserved pathways of RNA metabolism in all eukaryotes. Moreover, while the previously characterized Nab proteins have structural motifs common among metazoan hnRNPs, the overall structural arrangement of these motifs was dissimilar. Of particular note was the absence from this

small family of yeast hnRNPs of a representative of the most abundant metazoan hnRNPs, the A/B-type proteins. Finally, the relative inefficiency of the *in vivo* crosslinking procedure may have precluded the identification of hnRNPs that are weakly or only transiently associated with poly(A)$^+$ RNA or those hnRNPs of lower abundance.

With these considerations in mind, the first specific aim of my research project was to identify novel hnRNPs in yeast. Using a variation on the original isolation scheme, crosslinked RNPs were purified from a *nab2* mutant strain (Anderson, 1995) that exhibited elevated levels of poly(A)$^+$ RNA and Nab proteins in the nucleus. These observations suggested that these cells would be enriched for nuclear hnRNP-RNA complexes, which might help to overcome the inherent inefficiencies of *in vivo* crosslinking techniques. Indeed, this mutant proved to be a valuable biological reagent and facilitated the identification of a novel yeast hnRNP, Nab4p. This protein has remained the focus of my work.

DNA sequence data quickly revealed the successful identification of an hnRNP A/B-type protein, at least in regard to the overall structural organization of Nab4p. In order to gain clues about the potential functions of Nab4p, I undertook an extensive molecular characterization, including an analysis of the RNA-binding activities and subcellular distribution of Nab4p. To guide future genetic studies, I determined whether Nab4p was essential for cell viability. I also delineated the regions of Nab4p important for RNA-binding and subcellular localization. These studies revealed striking similarities between Nab4p and human hnRNP A/B proteins beyond the initial structural relatedness, suggesting that Nab4p may be a functional homologue of these metazoan hnRNPs.

At this point, my research turned to the functional analysis of Nab4p. To begin to identify the aspects of pre-mRNA processing and mRNA transport that are influenced by Nab4p, two general approaches were employed. First, conditional lethal *nab4* alleles were generated and used to alter the physiological concentration of functional Nab4p. The effect of these alterations on pre-mRNA processing and mRNA transport was evaluated. Second, proteins that interact with Nab4p or function in similar metabolic pathways were identified.

These functional studies suggested that Nab4p might influence a number of events during the biogenesis of mRNA in both the nucleus and cytoplasm of cells. Surprisingly, *nab4* mutants revealed a direct role for Nab4p in pre-mRNA 3'-end formation. The yeast *S. cerevisiae* provided an ideal experimental model to evaluate the function of Nab4p in 3'-end formation. Only a few yeast factors are required and sufficient to reconstitute 3'-end formation *in vitro*. Recently, Hrp1p, a protein identical to Nab4p, was identified as an essential component of one of these factors (Kessler et al., 1997). Therefore, this reconstitution system was utilized to investigate the contribution of Nab4p to 3'-end formation *in vitro*. The relevance of these findings to pre-mRNA processing *in vivo* was evaluated and will be discussed.

Isolation and Characterization of Nab4p, a Yeast hnRNP Similar to Metazoan hnRNP A/B Proteins

While investigating the function of the yeast hnRNP Nab2p, the temperature sensitive *nab2-13* allele was isolated and characterized (Anderson, 1995). Cells expressing mutant protein from this allele grew well at 30°C yet showed a striking accumulation of poly(A)$^+$ RNA in the nucleus, even during permissive growth. This

increase in poly(A)$^+$ RNA within the nucleus was concurrent with increases in the intranuclear concentrations of Npl3p/Nab1p and Nab3p (Anderson, 1995; A.M. Krecic and M.S. Swanson, unpublished data). From these observations, we reasoned that nuclear hnRNP-RNA complexes were enriched in the nucleus of *nab2-13* cells. Therefore, we attempted to isolate additional Nab proteins using this strain. Cells were grown at 30°C and subsequently irradiated with UV light to covalently crosslink RNA-binding proteins to their substrate RNAs. Crosslinked poly(A)$^+$ RNA-protein complexes were purified by oligo(dT)-cellulose chromatography and injected into mice to generate antibodies. The resulting polyclonal antisera were used to screen a λgt11 genomic expression library for the isolation of genes encoding novel Nab proteins. One of the partial genomic clones isolated encoded a novel protein that was named Nab4p because it satisfied the criteria previously established to authenticate these nuclear RNA-binding proteins as hnRNPs. First, monoclonal antibodies against Nab4p revealed a 70 kDa polypeptide in total yeast proteins that was also among those proteins crosslinked to poly(A)$^+$ RNA *in vivo* (Figure 3), while Pub2p, a ribosomal protein, was not. Second, Nab4p was localized predominantly to the nucleus of cells, as determined by indirect cellular immunofluorescence using anti-Nab4p monoclonal antibodies (Figure 4).

The full-length *NAB4* gene (Figure 5) was isolated by screening a YCp50 genomic DNA library by hybridization with a DNA fragment from the 5' end the λgt11 partial clone. The deduced amino acid sequence from *NAB4* revealed that Nab4p was an acidic protein (pI=5.4) of 534 amino acids with a predicted molecular weight of ~60 kDa. The primary structure of Nab4p contained several previously characterized RNA-binding motifs, including two RBDs in the amino-terminal half of the protein and a glycine-rich

Figure 3. Nab4p is associated with poly(A)$^+$ RNA *in vivo*. Polyclonal antisera raised against poly(A)$^+$ RNA-RNP complexes isolated from *nab2-13* cells was used to isolate a novel open reading frame, designated *NAB4*. Total cell extracts (total) and purified poly(A)$^+$ RNA-RNP complexes (crosslinked) from BJ926 cells were compared by immunoblot analysis using monoclonal antibodies (mAbs) 3H1 against Nab4p or 2B1 against Pub2p/Ssm1p, a large ribosomal subunit protein (Anderson et al., 1993b; Petitjean et al., 1995), respectively. Sizes are indicated in kilodaltons.

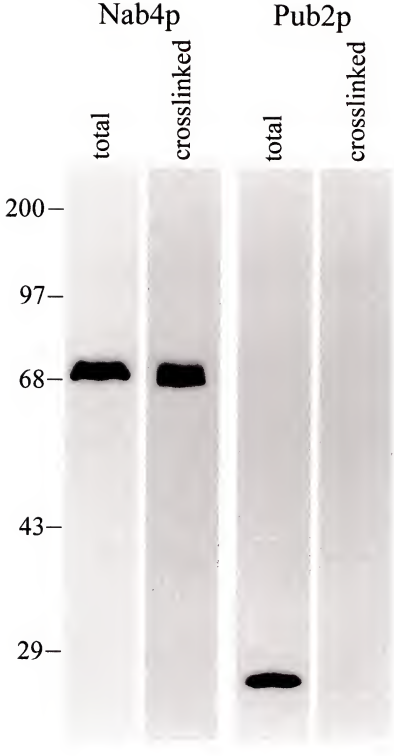

Nab4p
Pub2p
total
crosslinked
total
crosslinked
200–
97–
68–
43–
29–

Figure 4. Nab4p is a predominantly nuclear protein. The subcellular distributions of Nab4p and Pab1p were compared by indirect cellular immunofluorescence using mAbs 2D1 and 1G1, respectively. Antigens were localized in BJ926 cells using fluorescein- or rhodamine-conjugated secondary antibodies. Gray scale images are shown for Nab4p (a) and Pab1p (b). The DNA distribution is visualized by DAPI (c).

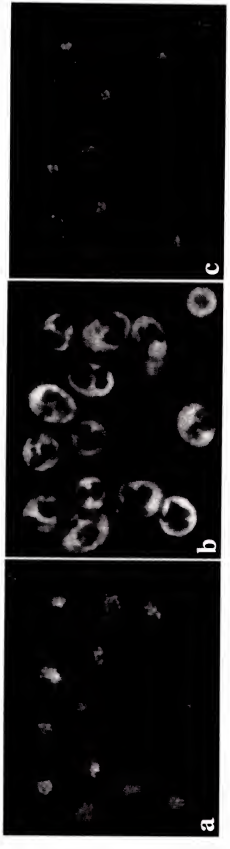
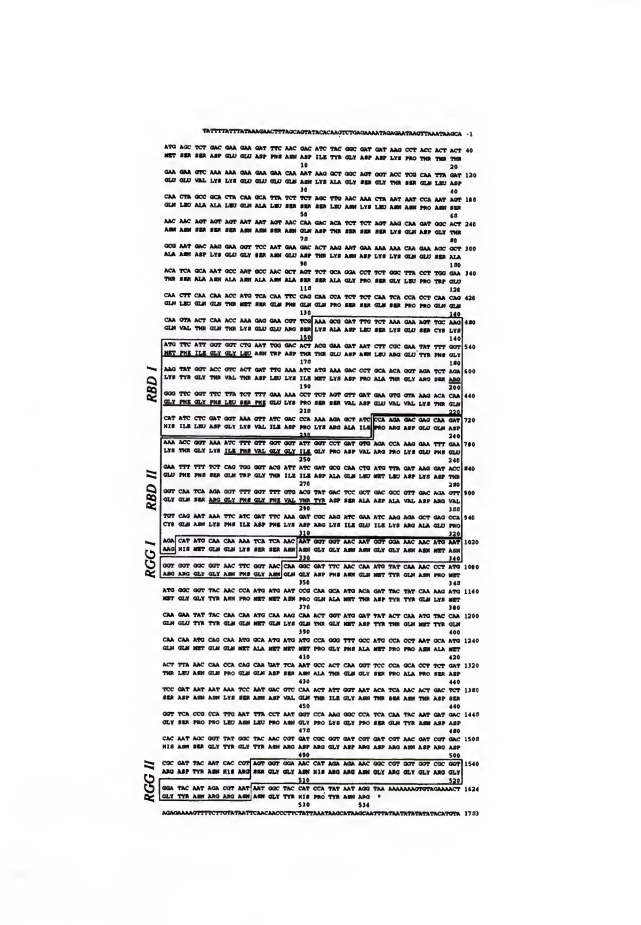

a
b
c

Figure 5. DNA and deduced amino acid sequence of *NAB4*. The genomic DNA sequence from nucleotide –69 to 1703 is shown. The large open boxes indicate the two ribonucleoprotein consensus sequence RNA-binding domains (RBD I and RBD II) with the underlined amino acids representing RNP1 and RNP2, conserved octapeptide and hexapeptide motifs, respectively, within RBDs. The two RGG boxes (RGG I and RGG II) are indicated by the smaller boxes. These sequence data are available from GenBank/EMBL/DDBJ under accession number U35737.


RBD I
RBD II
RGG I
RGG II

carboxyl-terminal half with two putative RGG box motifs. In addition, the amino terminus of Nab4p contained three regions rich in aspartate and glutamate (D/E), serine and asparagine (S/N), or glutamine and proline (Q/P). These regions are believed to be responsible for protein-protein interactions, and all of the Nab proteins described to date possess a Q/P-rich region. The carboxyl terminus of Nab4p also contained five repeats of the mixed-charge dipeptide aspartate and arginine (DR). This was similar to the alternating charge carboxyl-terminal domains of the U1 snRNA-associated 70 kDa protein (Theissen et al., 1996) and the 68 kDa subunit of mammalian cleavage factor I (Rüegsegger et al., 1996).

The array of RNA-binding motifs found in Nab4p was strikingly similar to the metazoan hnRNP A/B proteins (Figure 6). Because of this structural similarity, the RNA and single-stranded DNA (ssDNA) binding properties of Nab4p were compared to members of this group of hnRNPs (Figure 7A). Nab4p bound with preference for poly(G) and poly(U) RNA homopolymers, showing the same sequence preference as hnRNP A2, but considerably more Nab4p bound to poly(U) up to 0.5M NaCl. In addition, the chromatographic behaviors of Nab4p and hnRNP A1 on ssDNA-cellulose were similar, with both proteins eluting between 0.3-0.5 M NaCl (D.R. DeVore and M.S. Swanson, unpublished data). While the putative RGG box motifs in Nab4p were somewhat degenerate from the consensus motif, they were tested for RNA-binding capacity, independent of the full-length protein. Each was fused to the maltose binding protein and each fusion protein displayed the same binding preference to RNA homopolymers as full-length Nab4p (Figure 7B). Notably, the binding was significantly weaker, especially with the more degenerate RGG I. These similarities in primary

Figure 6. Nab4p is related to the yeast Npl3p/Nab1p and human hnRNP A1 and A2 proteins. (A) Diagrams of the primary structures of Nab4p, Npl3p/Nab1p (Bossie et al., 1992; Wilson et al., 1994), hnRNP A1 (Biamonti et al., 1994), and hnRNP A2 (Burd et al., 1989). The two different types of RNA-binding domains are indicated by solid (RBD) or stippled (RGG box) regions. (B) Sequence alignment of the two RBDs of Nab4p and the related RBDs from the human hnRNP D protein AUF1 (Ehrenman et al., 1994) and the neural specific *Drosophila* protein musashi (Nakamura et al., 1994). (C) Sequence alignment of the two RGG boxes of Nab4p with Nab2p (Anderson et al., 1993b), Npl3p/Nab1p, and the human hnRNP A1, A2, and U proteins (Kiledjian and Dreyfuss, 1992).

A

B

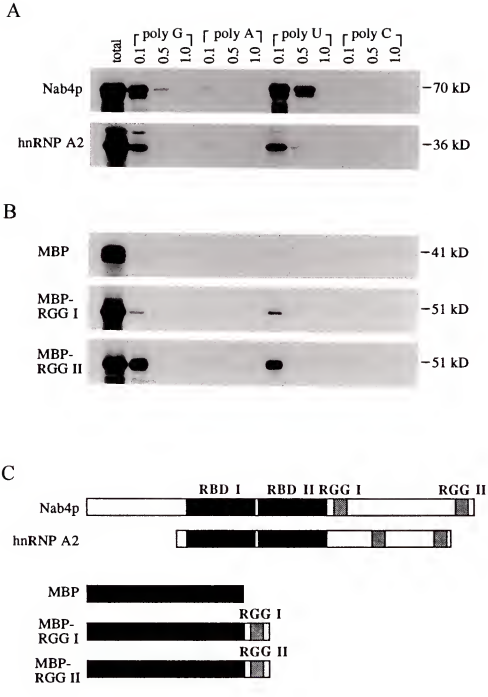
```
Nab4 RBDI   ESCK MFIGGL NWDTTEDNLREVFGKYGVTVTDLKIMKDPATGRS RGFGFLEF EKPSSVDEVVKTQ-HILDGKVIDPKRAIPR
AUP1        DEGK MFIGGL SWDTTKKDLKDYFSKFGEVVDCTLKLDPITGRS RGFGFVLF KEESSVDKVMDQKEHKLNGKVIDPKRAEAM

Nab4 RBDII  KTGK IPVGGI GPDVRKEEPEPPSQWGTIIDAQLMLDKDTQGS RGFGFVTY DSADAVDRVCQNKFIDFKDRKIEIKRAEPR
            KTKK IPVGGV SQDTSAEEVKAYFSQFGPVEETVMLMDQQTKRH RGFGFVTF ENEDVVDRVCEIHFHTIKMEKVECKKAQPK

Consensus       LFVGNL      E  L  F  FG I              K  RGFGFV F          A          L G
                IYIKG       D           Y V              R  R YA  Y                     I
```

C

Protein

Nab2p	101		RGG	GAVG	KNR		RGG		RGG	N	RGG	200
Nab4p	331		NGG	N	NGG		NNM	NR	RGG	N	PGN	349
	508		SGG	WNR	RNG		RGG		RGG	YN	RRN	527
Npl3p/												
Nab1p	287		RGG		FRG		RGG	F	RGG	F	RGG	307
	310		RGG		FGG		RGG		FGG	P	RGG	316
	327		YGG	Y S	RGG		YGG		RGG		YGG	345
	347		RGG	YDSP	RGG	YDSP	RGG	YS	RGG		YGG	371
human												
hnRNPA1	194		RGR	SG	RGN		FGG	G	RGG	G	FGG	212
	218		RGG	NFSG	RGG		FGG	S	RGG	GG	YGG	239
human												
hnRNPA2	186		RSR	SG	RGG	NFG	FGD	S	RGG	GG	MFG	208
	239		PGG	GN	FGG	SFG	YGG	G	RGG		YGG	259
human												
hnRNPU	696		RGG	NF	RGG		APG	N	RGG	YN	RRG	715

Figure 7. Nab4p exhibits the same sequence-preference binding to RNA as hnRNP A/B proteins. (A) Recombinant Nab4p and human hnRNP A2 proteins were synthesized by *in vitro* transcription/translation as described in Materials and Methods. ^{35}S-methionine-labeled proteins were tested for the ability to bind to poly(G), poly(A), poly(U) and poly(C) covalently bound to agarose beads at 0.1-1.0 M NaCl. (B) RNA-binding activity of RGG box-maltose-binding protein (MBP) fusion proteins transcribed and translated from pNAB4.13 (MBP-RGG I) and pNAB4.14 (MBP-RGG II) (see Appendix). The binding of the MBP alone is compared to MBP-RGG I and MBP-RGG II fusion proteins. Sizes are indicated in kilodaltons. (C) Schematic representation of Nab4p, human hnRNP A2 and the MBP fusions used in these experiments.


A
total
poly G
0.1
0.5
1.0
poly A
0.1
0.5
1.0
poly U
0.1
0.5
1.0
poly C
0.1
0.5
1.0
Nab4p
−70 kD
hnRNP A2
−36 kD
B
MBP
−41 kD
MBP-
RGG I
−51 kD
MBP-
RGG II
−51 kD
C
RBD I
RBD II
RGG I
RGG II
Nab4p
hnRNP A2
MBP
MBP-
RGG I
RGG I
MBP-
RGG II
RGG II

structure and nucleic acid binding activity between Nab4p and hnRNP A1/A2 suggested that Nab4p may indeed be a homologue of the metazoan A/B proteins.

Cytoplasmic Accumulation of a Yeast Nucleocytoplasmic Shuttling hnRNP

Another distinctive feature of the human hnRNP A/B proteins is the presence of a transport signal, termed M9. This sequence mediates both nuclear import and nuclear export of these proteins as well as nucleocytoplasmic mRNA export (Siomi and Dreyfuss, 1995; Michael et al., 1995; Pollard et al., 1996; Izaurralde et al., 1997). While there was no obvious sequence in Nab4p related to the M9 signal, a functionally related signal might exist. To delineate the region of Nab4p required for nuclear localization of Nab4p, an in-frame fusion protein was generated between a green fluorescent protein (GFP) and Nab4p (GFP-Nab4p, Figure 8). This fusion protein was localized primarily in the nucleus of cells (Figure 9A) and complemented a *nab4* null allele (A.M. Krecic and M.S. Swanson, unpublished data). Deletion of the carboxyl-terminal 87 amino acids of Nab4p, including one of the RGG box motifs, led to mislocalization of the fusion protein to the cytoplasm. This was also true of GFP-Nab4p fusion proteins truncated progressively further from the carboxyl terminus (Figure 8), none of which complemented a *NAB4* deletion. To determine if this region alone could mediate nuclear localization, portions of the carboxyl terminus were fused to GFP and their localization in cells was monitored. The carboxyl-terminal 47 amino acids of Nab4p, containing RGG II and the mixed-charge DR dipeptide region (see Figure 5), were sufficient for nuclear localization of GFP (Figure 9, NLS) and it was subsequently shown that the DR repeats were expendable for this activity (A.M. Krecic and M.S. Swanson, unpublished data).

Figure 8. Efficient nuclear localization of Nab4p requires the carboxyl terminus. A GFP-*NAB4* fusion construct was generated and truncated progressively from the carboxyl terminus of Nab4p as described in Materials and Methods. GFP-Nab4p fusion proteins were expressed in YAO403 and their steady-state distributions were monitored by fluorescence microscopy. The distribution of truncated fusion proteins between the nuclear and cytoplasmic compartments is indicated relative to full-length GFP-Nab4p. The GFP-*NAB4* constructs used were, from top to bottom, pNAB4.81, pNAB4.81a, pNAB4.81b, pNAB4.81c, pNAB4.81d, pNAB4.81e, pNAB4.81f, and pNAB4.81g.

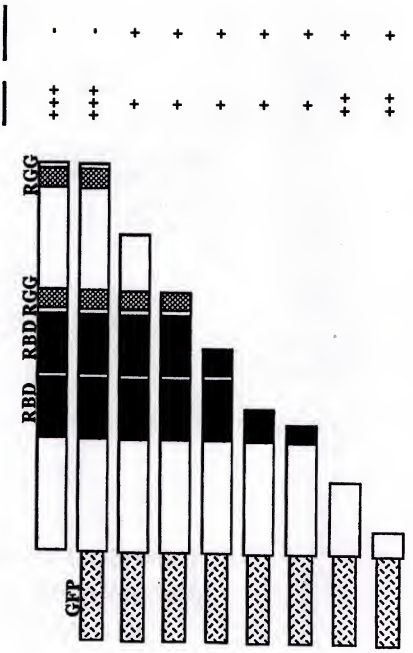
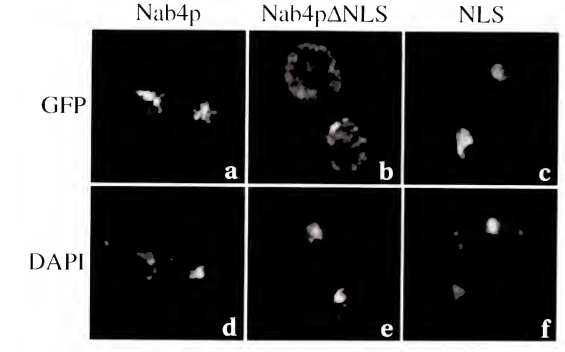

NUC
CYTO
+++
+++
+
+
+
+
+
+
+
+
+
+
+
+
RGG
RBD
RBD
RGG
RBD
GFP

Figure 9. The carboxyl terminus of Nab4p is necessary and sufficient for nuclear localization. The distribution of various GFP-Nab4p fusion proteins was analyzed by digital fluorescence microscopy. Images are shown for YAO403 carrying (a) full-length GFP-Nab4p (pNAB4.76), (b) pNAB4.81a, and (c) pNAB4.88, a fusion between GFP and the 47 carboxyl terminal amino acids of Nab4p containing RGG II and the DR dipeptide repeats. Fluorescence images are compared to DAPI staining for DNA (d-f). Shown below is the sequence alignment of the M9 NLS of human hnRNP A1 and A2 proteins (Pollard et al., 1996) and yeast Nab4p. Also shown is a related sequence from the *Drosophila* hnRNP A/B protein, hrp36 (Visa et al., 1996a). The primary structures of hnRNP A1 and Nab4p are schematized below the sequence alignment. The RNA-binding domains (RBD), nuclear localization sequences (M9, NLS) and the glycine (Gly) or glycine/glutamine-rich (Gly/Gln) regions are indicated.



Human A2	296	Q[illegible]NY[illegible]S[illegible]NMG[illegible]N-[illegible]GPGGSGGS[illegible]
Human A1	268	N[illegible]S-[illegible]SS-[illegible]Q[illegible]FAKP[illegible]Q-[illegible]
Yeast Nab4	503	[illegible]---[illegible]--NR[illegible]NN--[illegible]YHPYNR
Fly hrp36	276	[illegible]-GG-[illegible]---[illegible]GGGNS[illegible]GSW[illegible]GNGGGG

Alternatively, because of the relatively small size (~30 kDa), these GFP-Nab4p NLS fusion proteins may enter the nucleus by diffusion through NPCs. That these fusion proteins accumulate in the nucleus may be due to a retention mechanism. By analogy with metazoan NRS sequences, this carboxy-terminal region of Nab4p may exhibit a similar activity, or this region of Nab4p may interact with a nuclear component that is actively retained in the nucleus. However, this latter explanation for nuclear accumulation is less likely for the following reasons. Recently, Kessler et al. reported that Hrp1p, identical to Nab4p, has the ability to shuttle between the nucleus and cytoplasm (Kessler et al., 1997), another property common among hnRNP A/B proteins. Although the nucleocytoplasmic shuttling hnRNP proteins contain sequences that possess both NLS and NES activities, NRS sequences have not been identified in these proteins. In contrast, the NRS-containing hnRNP C proteins are not nucleocytoplasmic shuttling proteins, suggesting that the presence of a NRS would be incompatible with shuttling activity. There may be a NRS in the amino terminus of Nab4p, which could explain the modest nuclear accumulation observed for the two shortest GFP-Nab4p fusion protein, containing only the most amino-terminal portions of Nab4p (refer to Figure 8 legend, pNAB4.81f and pNAB4.81g). Beyond both being glycine-rich, little sequence identity exists between the M9 signal and the Nab4p NLS defined here (Figure 9). However, the position of the Nab4p NLS near the carboxyl terminus is similar to hnRNP A/B proteins, and the *Drosophila* hnRNP A/B protein hrp36 contains a closely related 36 amino acid sequence. Therefore, on the basis of both structural and functional similarities, Nab4p is clearly most related to the metazoan hnRNP A/B proteins, although it remains to be determined whether the Nab4p NLS also functions in nuclear export of Nab4p or mRNA.

While Kessler et al. demonstrated that Nab4p/Hrp1p was a nucleocytoplasmic shuttling protein, several key observations had been made during my studies that highlight the importance of this activity in the context of mRNA nucleocytoplasmic transport. First, Nab4p interacted with Nab2p, another nucleocytoplasmic shuttling protein (Truant et al., 1998), which has been implicated in mRNA export (Anderson, 1995). This interaction was demonstrated by employing the two-hybrid system (Figure 10A), commonly used to identify proteins that interact *in vivo*. While the interaction between Nab4p and Nab2p was the strongest observed in this screen (Figure 10B), two additional proteins, Gbp2p and Mgm101p, were found to interact with Nab4p (Figure 10, A and B). The possible significance of these interactions will be discussed below. The two-hybrid interaction between Nab4p and Nab2p was corroborated by co-immunoprecipitation of Nab4p and Nab2p from cells labeled *in vivo* (M.S. Swanson, personal communication). Second, Blobel and coworkers reported that Nab2p and Nab4p were co-isolated by affinity chromatography with Kap104p, a cytosolic protein that binds to nuclear pore complex proteins (Aitchison et al., 1996). Finally, and one of the most intriguing observations made during these studies, the subcellular distribution of Nab4p changes depending on the growth conditions. Under certain growth conditions requiring rapid bursts of transcriptional activity from subsets of nuclear-encoded genes, Nab4p accumulated in the cytoplasm of cells. When cells were grown continuously in YPD, which represses the transcription of nuclear-encoded mitochondrial genes, the distribution of Nab4p was predominantly nuclear, as indicated by comparison with the DNA-staining pattern (Figure 11, a and b). In contrast, when cells were grown in derepressing SSL media, a marked cytoplasmic accumulation of Nab4p was observed

Figure 10. Identification of proteins interacting with Nab4p *in vivo*. (A) Summary of proteins identified via the two-hybrid screen. (B) Quantitative analysis of interactions with Nab4p. A full-length representative of each class of cDNA clones isolated was used to assay the interaction with Nab4p *in vivo*. (The *MGM101* cDNA clone lacked the DNA sequence encoding the first 7 amino acids of Mgm101p.) Interactions were measured by a quantitative liquid assay and are expressed as β–gal specific activity. Values were averaged from samples assayed in triplicate from two separate experiments and standard deviations are indicated.

A

In vivo interactions identified by the two hybrid system. Indicated are the genes corresponding to the clones isolated from a yeast cDNA library using *NAB4* as bait.

Chromosomal locus	Independent isolates*	Function
NAB2	4	hnRNP; influences poly-adenylation and mRNA export; interacts with human Rab
GBP2	6	single-stranded G-strand telomeric DNA-binding protein
MGM101	1	required for mitochondrial genome maintenance

*Out of approximately 250,000 Trp^+ Leu^+ transformants, 27 His^+ colonies were screened by filter assay for ß-gal activity. Of these, 22 were ß-gal^+. Self-activating and false-positive clones were eliminated and the remaining 11 were identified by DNA sequencing.

B

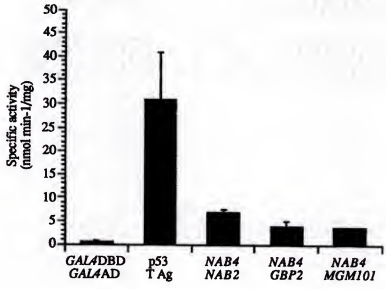
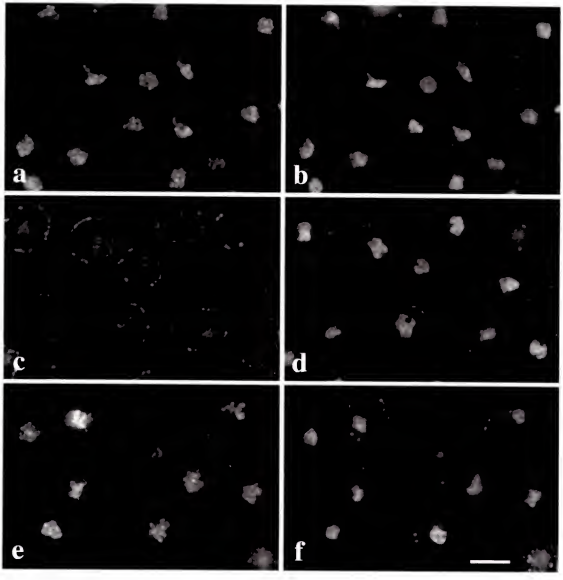


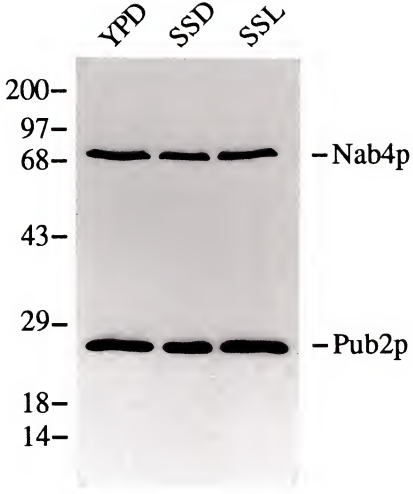
Figure 11. The subcellular distribution of Nab4p varies depending upon whether cells are grown on fermentable or non-fermentable carbon sources. The distributions of Nab4p and Nab2p were determined by indirect cellular immunofluorescence and three-dimensional optical microscopy. Wild-type cells were grown in YPD (a and b) or SSL (c to f). Nab4p was detected using mAb 3H1 (a and c), Nab2p with mAb 3F2 (e) and DNA with DAPI (b, d, and f). Bar=4 μm.


a
b
c
d
e
f

with a concomitant ~50% decrease in the apparent nuclear concentration of Nab4p (Figure 11, c and d). However, immunoblot analysis indicated no significant change in the steady-state intracellular concentration of Nab4p (Figure 12). In addition, Nab4p was not distributed uniformly throughout the cytoplasm, but was mainly punctate, with a subpopulation of the protein being present near the periphery of cells in the vicinity of mitochondria. To study how rapidly Nab4p would appear in the cytoplasm following shift from a fermentable to a non-fermentable carbon source, cells were grown in YPD and shifted into SSL following centrifugation and washing. The punctate appearance of Nab4p in the cytoplasm was rapid, occurring within 15 minutes after the shift into SSL (A.M. Krecic and M.S. Swanson, unpublished data). The cytoplasmic distribution of Nab4p was also compared to the distribution of the β subunit of the F1 ATPase, an inner mitochondrial membrane protein, and a subpopulation of cytoplasmic Nab4p appeared to co-localize with this mitochondrial protein (A.M. Krecic and M.S. Swanson, unpublished data). Other yeast hnRNPs, including Nab2p, did not exhibit this cytoplasmic accumulation (Figure 11, e and f).

Does Nab4p play a role in mitochondrial biogenesis? According to the results of the two-hybrid screen described above, Mgm101p, a nuclear-encoded mitochondrial protein required for maintenance of the mitochondrial genome (Chen et al., 1993), can interact with Nab4p *in vivo* (Figure 10). This interaction was relatively weak and *MGM101* was rescued just once during this screen, suggesting that it may not represent a physiologically significant interaction. However, this observation and the re-distribution of Nab4p during mitochondrial induction suggested that Nab4p might be important for mitochondrial biogenesis. Since Nab4p appeared to accumulate in the cytoplasm in a

Figure 12. Immunoblot analysis of total proteins isolated from cells grown in YPD, semi-synthetic glucose (SSD) and SSL. Total proteins were isolated from cells grown to mid-log phase and blots were probed with mAb 3H1 and 2B1 against Nab4p and Pub2p, respectively. Sizes are indicated in kilodaltons.


YPD
SSD
SSL
200–
97–
68–
43–
29–
18–
14–
–Nab4p
–Pub2p

peri-mitochondrial distribution pattern, mitochondrial isolation studies were performed to determine whether Nab4p could be co-purified with mitochondria. When cells were grown in SSL to induce mitochondrial biogenesis, immunoblot analysis of a purified mitochondrial fraction revealed the presence of Nab4p (A.M. Krecic and M.S. Swanson, unpublished data). However, Nab2p and the major cytoplasmic polyadenylate tail binding protein Pab1p also co-purified with mitochondrial fractions. Therefore, it was not possible to distinguish between Nab4p that was associated with mitochondria *in vivo* versus Nab4p that adventitiously associated during subcellular fractionation *in vitro*. Based on these results and the structural similarities between Nab4p and both the hnRNP A/B proteins and the AUF1/hnRNP D proteins, I speculated that Nab4p plays a role in the cytoplasm of cells, perhaps in mRNA trafficking or stability. However, the relevance of these observations to the function of Nab4p remains to be investigated.

The Intracellular Level of Nab4p Is Critical for Normal Cell Growth

Since the metazoan hnRNP A/B proteins have been implicated in the regulation of pre-mRNA processing, the function of Nab4p *in vivo* was investigated using a genetic approach. To do this in yeast, it was first important to determine whether *NAB4* was an essential or non-essential gene. Genomic DNA blot analysis indicated that *NAB4* was a single-copy gene and RNA blot analysis revealed that *NAB4* was expressed as a single polyadenylated mRNA of approximately 1800 nucleotides (Figure 13). To determine whether *NAB4* was essential for cell viability, one chromosomal allele in a wild type diploid was replaced using a subclone of *NAB4* in which 50% of the open reading frame (ORF) had been replaced by the selectable marker *LEU2* (Figure 14A). Genomic DNA

Figure 13. *NAB4* is a single-copy gene. (A) Chromosomal locus of *NAB4*. The restriction map includes the genomic region containing the entire *NAB4* gene and the 3'-end of *SMF1* (West et al., 1992). Open reading frames are shown as rectangles below the restriction map and arrowheads indicate the direction of transcription. The bracketed EcoRI site indicates the position of a synthetic linker present on the λgt11 clone pRNP4.0. (B) Southern blot analysis of BJ926 genomic DNA digested with multiple restriction endonucleases, demonstrating that *NAB4* is a single-copy gene. Sizes are indicated in kilobases. (C) Northern blot analysis of poly(A)$^+$ RNA from BJ926 revealed one major transcript of 2.2 kb encoded by *NAB4*. The DNA probe used in these experiments was the 330 bp SstI-PstI fragment located near the 5' end of the *NAB4* coding region.

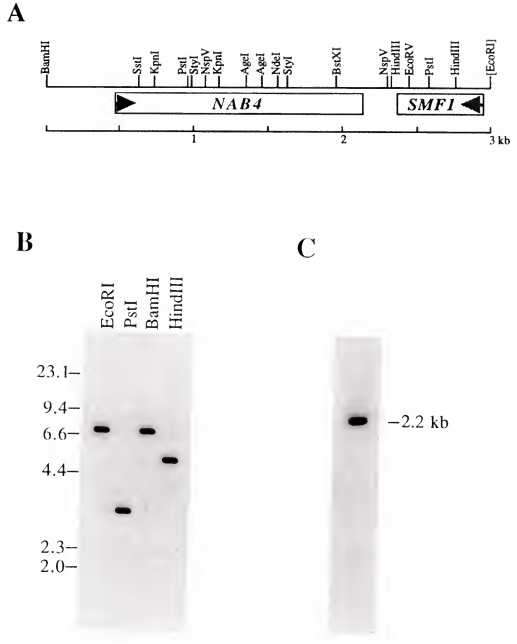
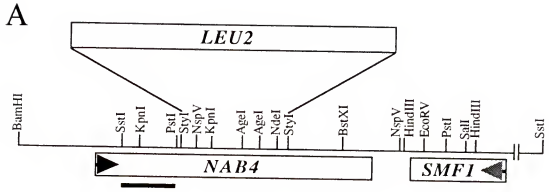
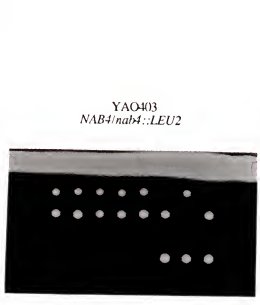

A
BamHI
SstI
KpnI
PstI
StyI
NspV
KpnI
AgeI
AgeI
NdeI
StyI
BstXI
NspV
HindIII
EcoRV
PstI
HindIII
EcoRI
NAB4
SMF1
1
2
3 kb
B
EcoRI
PstI
BamHI
HindIII
23.1–
9.4–
6.6–
4.4–
2.3–
2.0–
C
–2.2 kb

Figure 14. *NAB4* is essential for cell viability. (A) One allele of *NAB4* was disrupted by replacing ~40% of the open reading frame with *LEU2* by targeted single-gene transplacement. (B) Southern analysis of genomic DNA from a Leu^+ diploid, YAO403, was performed to confirm recombination within the *NAB4* locus (note the appearance of the 7.1 kb recombinant fragment in addition to the wildtype 5.5 kb fragment). (C) Segregation analysis of YAO403 gave a 2:2 pattern of viable to non-viable cells, indicating that *NAB4* is essential for growth.


A
LEU2
BamHI
SstI
KpnI
PstI
StyI
NspV
KpnI
AgeI
AgeI
NdeI
StyI
BstXI
NspV
HindIII
EcoRV
PstI
SalI
HindIII
SstI
NAB4
SMF1


B
YJA501
YAO403
7.1 kb
5.5 kb


C
YAO403
NAB4/nab4::LEU2

blot analysis confirmed the desired recombination event (Figure 14B) and segregation analysis of a leucine prototroph (Leu^+) gave a 2:2 pattern of viable to non-viable spores (Figure 14C). All viable spores were Leu^-, indicating that, like the previously described *NAB* genes, *NAB4* is essential for cell viability. Leu^+ haploid spores could be rescued using a 2.7 kb BamHI/SalI genomic fragment from the YCp50 clone encoding *NAB4* (see Figure 14A), indicating that only the *NAB4* gene had been disrupted in these cells. The *NAB4* gene was located on chromosome XV and did not appear to contain any intron sequences.

Our model for the function of hnRNPs in the regulation of gene expression predicts that changing the intracellular concentration of hnRNPs might alter the protein-RNA complexes in which they exist and thereby affect mRNA biogenesis and transport. To study the effect on cell growth of altering the concentration of Nab4p, *NAB4* was placed under the control of the inducible/repressible *GAL1* promoter. When YAO407, a diploid cell harboring this construct (*NAB4/nab4::LEU2*), was sporulated and dissected on rich media containing glucose (YPD) or galactose (YPGal), haploid cells carrying the *nab4* null allele and the *GAL::NAB4* plasmid could not be recovered. In addition to confirming that *NAB4* was essential for viability, this experiment indicated that overexpression of Nab4p was inhibitory to cell growth. Consistent with this was the recovery of haploid cells containing both the *NAB4* chromosomal allele and the *GAL::NAB4* allele when YAO407 was dissected on minimal media containing raffinose (SR), a non-inducing/non-repressing carbon source. These cells were not dependent on the *GAL::NAB4* plasmid for growth and haploid cells disrupted at *NAB4* could not be recovered under these conditions. These results precluded using this method to study the

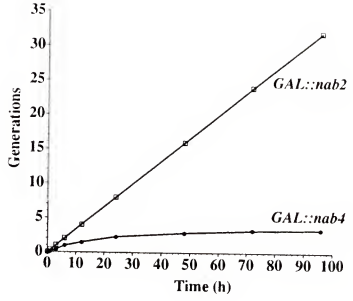
effect of depleting Nab4p from cells. To confirm that overexpression inhibited growth, YAO407 cells were grown in SR and monitored for growth following the addition of galactose to cultures to induce expression of Nab4p (Figure 15). Growth was severely inhibited within 1 h even though Nab4p expression had increased only 2-3 fold. Cells began to exhibit abnormal nuclear morphology as Nab4p continued to accumulate in the nucleus. This accumulation persisted over time and nuclear integrity was lost as Nab4p steady-state levels increased 10-20 fold. At various times after induction, cells were returned to a repressing carbon source and there was a direct correlation between the time following the addition of galactose and loss of cell viability. By 6 h post-induction, a time at which Nab4p levels had increased about 10-fold, less than 10% of the cells remained viable, indicating that increased levels of Nab4p are toxic to cells (Table 1). This was corroborated by several observations. First, when *NAB4* was provided on a low copy *CEN* plasmid (2-4 copies/cell), haploid cells were viable at 24°C and 30°C whether they contained a null or wild type allele of *NAB4*. However, in the presence of the wild type allele, cells grew noticeably slower and were

Table 1. Effect of Nab4p overexpression on cell viability.

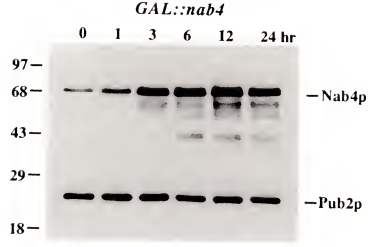
Time Post-Induction (h)	Cell Viability (%)
0	100
3	78
6	7
12	2

Figure 15. Overexpression of Nab4p is toxic to cells. (A) Overexpression of Nab4p rapidly inhibits cell growth. The number of generations that cells undergo while expressing Nab2p or Nab4p under the regulation of the *GAL1* promoter following the addition of 2% galactose at time 0 to cells growing in SR. Cells were kept in mid-log phase by the addition of SG as needed. (B) Immunoblot anaylsis of cells grown as in (A). The Nab4 and Pub2 proteins were detected with mAbs 3H1 and 2B1, respectively. Sizes are indicated in kilodaltons.

A


35
30
25
20
15
10
5
0
Generations
GAL::nab2
GAL::nab4
0 10 20 30 40 50 60 70 80 90 100
Time (h)

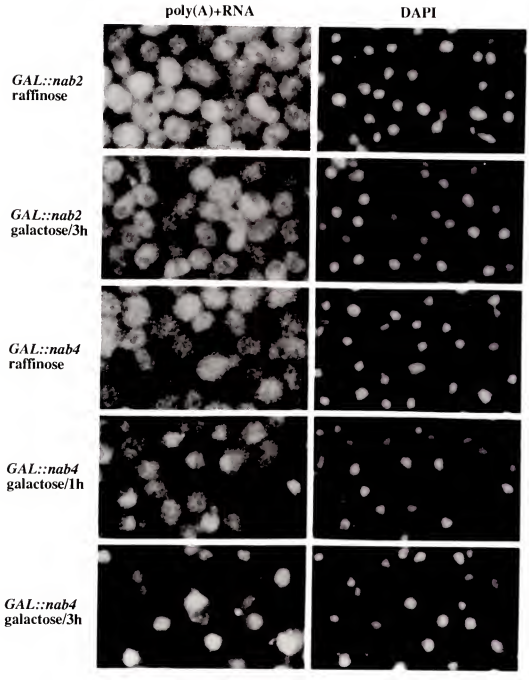
B


GAL::nab4
0 1 3 6 12 24 hr
97
68
43
29
18
Nab4p
Pub2p

severely growth inhibited at 36°C. Finally, initial attempts to use *NAB4* as bait in the two-hybrid system, in which *NAB4* was provided on a high copy plasmid, failed. Transformants containing this plasmid only produced pin-point colonies which were not viable. However, this phenomenon appeared to be dependent on the genetic background, as the two-hybrid system was successful using a different host strain.

Why, unlike the other Nab proteins, must the steady-state concentration of Nab4p be so carefully regulated *in vivo*? To address this question, the effect of overexpression of Nab4p on mRNA biogenesis and transport was examined. Since the human hnRNP A1 protein shuttles between the nucleus and cytoplasm, a model has been proposed that nucleocytoplasmic mRNA export requires the formation of a specific hnRNP-mRNA complex. This model predicts that altering the intracellular concentration of the A1 protein might alter the structure of the hnRNP-mRNA complex and affect mRNA export. Because Nab4p and hnRNP A1 appeared to be closely related, functional studies on Nab4p were continued by testing the model that *NAB4* expression is required for nuclear mRNA export. At various times after the addition of galactose, YAO407 cells were examined by in situ hybridization with oligo(dT)$_{50}$ for the subcellular distribution of mRNA (Figure 16). Cells exhibited a dramatic accumulation of poly(A)$^+$ RNA in the nucleus within 1 h after induction, with approximately 50% of all cells exhibiting this phenotype by 3 h after the addition of galactose. This accumulation persisted over time suggesting a block to mRNA export. However, rather than a block to export, this phenotype could arise indirectly due to defects in pre-mRNA splicing or 3'-end formation. While the metazoan hnRNP A/B proteins have been implicated in regulating splice site selection in a dose-dependent fashion, only 2-5% of all yeast genes contain

Figure 16. Cells overexpressing Nab4p accumulate poly(A)$^+$ RNA in the nucleus. The subcellular distribution of poly(A)$^+$ RNA was determined by in situ hybridization with oligo(dT)-digoxigenin. Cells expressing either Nab2p or Nab4p as in Figure 15 are shown grown in SR and at 1 h or 3 h after the addition of 2% galactose.

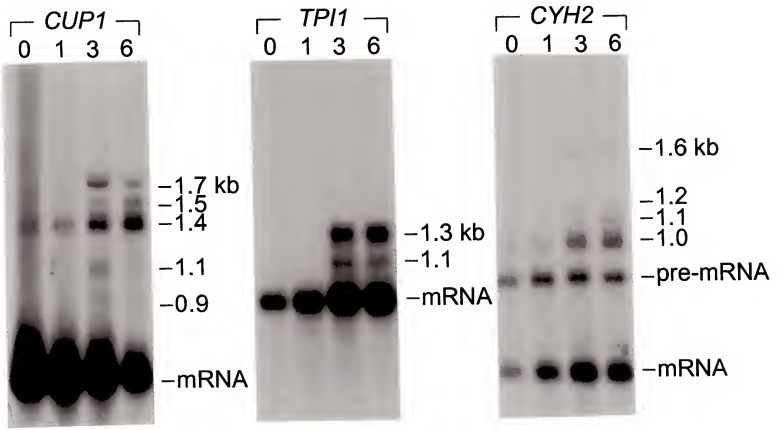

poly(A)+RNA
DAPI
GAL::nab2
raffinose
GAL::nab2
galactose/3h
GAL::nab4
raffinose
GAL::nab4
galactose/1h
GAL::nab4
galactose/3h

intron sequences. Therefore, it was questionable whether a block in pre-mRNA splicing could account for the profound nuclear accumulation of poly(A)$^+$ RNA observed. Indeed, when the steady-state level of pre-mRNAs and mRNAs from *ACT1* and *CYH2* were compared at various times after Nab4p induction, there was no change in the ratio of pre-mRNA to mRNA (Figure 17; A.M. Krecic and M.S. Swanson, unpublished data). Surprisingly, there was a significant accumulation of aberrantly long transcripts from a variety of genes. Interestingly, mutations in genes involved in mRNA export and pre-mRNA 3'-end processing have been shown to cause significant alterations in the length of both *CUP1* and *CYH2* transcripts (Forrester et al., 1992; Chanfreau et al., 1996). Therefore, these observations are consistent with a defect in mRNA export or pre-mRNA 3'-end processing, or both. A major limitation of these studies is that during overproduction of a protein *in vivo*, it may interact with and sequester a second protein. This could effectively inhibit the normal activity of the second protein and lead to indirect or pleiotropic effects. Because of the multiple phenotypes seen during overexpression of Nab4p, it was possible that any one or a combination of the defects suggested above might contribute to the cessation of cell growth. To distinguish between these possibilities, conditional-lethal *nab4* mutants were created.

Isolation and Characterization of *nab4* Mutants

Mutant *nab4* alleles were generated by random PCR mutagenesis and screened for a temperature sensitive (Ts$^-$) growth phenotype at 36°C. Even modest increases in Nab4p levels could cause, or contribute to, growth inhibition, so putative Ts$^-$ alleles were subcloned into the single-copy YCp50 plasmid to confirm the Ts$^-$ phenotype in the

Figure 17. Overexpression of Nab4p leads to the accumulation of aberrantly long transcripts for three genes. Total RNA was isolated from YAO407 cells grown in SR or at various times after the addition of 2% galactose. Number of hours post-induction are indicated by numerals above lanes and transcript sizes are indicated in kilobases. The normal mRNAs, and pre-mRNA for *CYH2*, are indicated.


CUP1
0 1 3 6
–1.7 kb
–1.5
–1.4
–1.1
–0.9
–mRNA
TPI1
0 1 3 6
–1.3 kb
–1.1
–mRNA
CYH2
0 1 3 6
–1.6 kb
–1.2
–1.1
–1.0
–pre-mRNA
–mRNA

presence of a *nab4* null allele. Three recessive Ts^- alleles were isolated, *nab4-1*, *nab4-4*, and *nab4-7*, each containing several missense mutations (Figure 18). Strains harboring each of these mutant alleles were compared to wild type for growth at 24°C and 36°C Figure 19A). In addition, growth in culture showed that these Ts^- mutations caused a rapid growth inhibition within 1-2 h after a shift to the non-permissive temperature (Figure 19B).

Knowing that the intracellular level of Nab4p is critical for normal cell growth, the level of mutant Nab4 proteins in these strains was determined by immunoblot analysis. At 24°C, each of the *nab4* mutants produced steady-state levels of Nab4 protein that were 2-3 fold higher than wild type Nab4p and stable even after a 2h shift to 36°C (Figure 19C). The elevated levels of mutant proteins were not due to episomal expression, since similar observations were made when comparing the expression of *NAB4* and *nab4-1* from plasmids of different copy numbers or following integration of *nab4-1* into the *NAB4* chromosomal locus (Figure 20). The production of Nab4p at 36°C indicated that the loss of cell viability was not due to a loss of Nab4 protein. However, the conditional loss of cell viability in *nab4* mutants was not due solely to the elevated levels of Nab4 protein in these cells, since these elevated levels were observed in mutants even at the permissive temperature. Furthermore, the increased steady-state concentration of mutant proteins did not lead to a gain of function, since each of these mutants was recessive to *NAB4*. These preliminary observations made using *nab4* mutants emphasized the importance of tightly regulating the level of Nab4p in cells and suggested that Nab4p may possess a functional activity that is regulated by the concentration of Nab4 protein. More importantly, this genetic screen provided several

Figure 18. Schematic representation of *nab4* Ts$^-$ mutant proteins. All *nab4* alleles were sequenced and the arrows and single letter abbreviations indicate the deduced amino acid substitutions as compared to wildtype Nab4p.

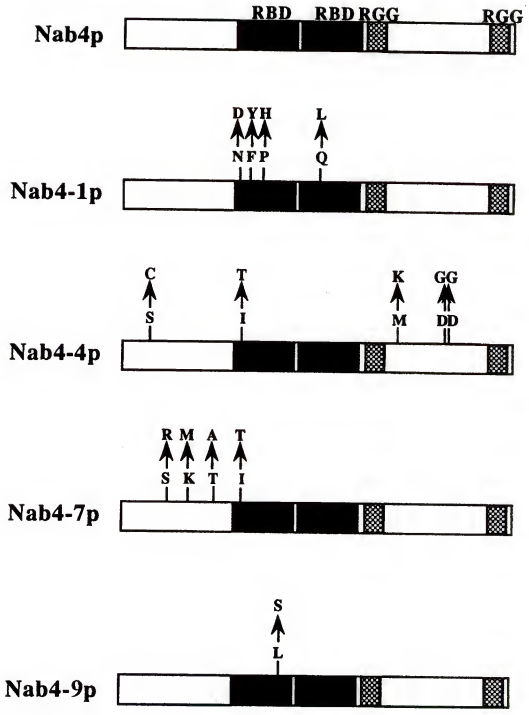

Nab4p
RBD
RBD
RGG
RGG
Nab4-1p
D
Y
H
L
N
F
P
Q
Nab4-4p
C
S
T
I
K
M
GG
DD
Nab4-7p
R
M
A
T
S
K
T
I
Nab4-9p
S
L

Figure 19. Characterization of recessive Ts$^-$ *nab4* alleles. (A) Growth of mutant strains *nab4-1* (YAO418-1D), *nab4-4* (YAO428-1B), and *nab4-7* (YAO431-1C) are compared to wildtype (YAO417-1D) at 24°C and 36°C. (B) Generation curve of *nab4* mutants. Cells were grown in YPD at 24°C and then shifted to 36°C. Growth rates are represented as the number of generations which cells undergo over time (*NAB4*, closed squares; *nab4-1*, closed circles; *nab4-4*, closed triangles; *nab4-7*, closed diamonds). (C) Mutant Nab4 proteins are stable at the non-permissive temperature. Total protein was isolated from cells grown in YPD at 24°C or shifted to 36°C for 2 h. Immunoblot analysis was performed using mAbs 3H1 and 2B1 against Nab4p and Pub2p, respectively. Sizes are indicated in kilodaltons.

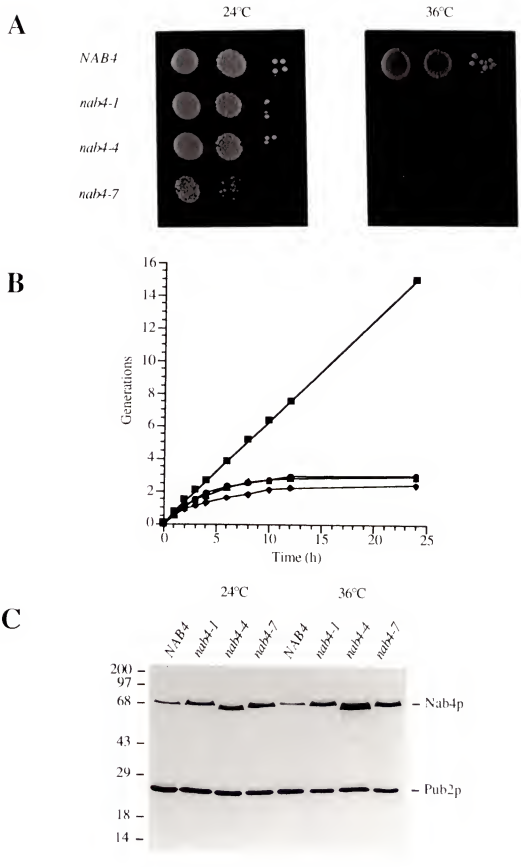
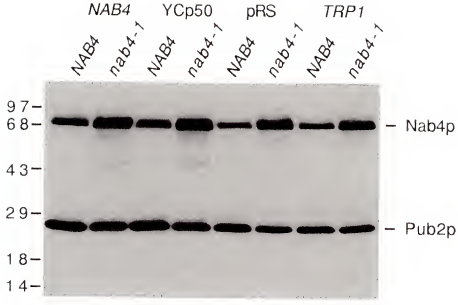

A
24°C
36°C
NAB4
nab4-1
nab4-4
nab4-7
B
Generations
Time (h)
C
24°C
36°C
NAB4
nab4-1
nab4-4
nab4-7
NAB4
nab4-1
nab4-4
nab4-7
200
97
68
43
29
18
14
Nab4p
Pub2p

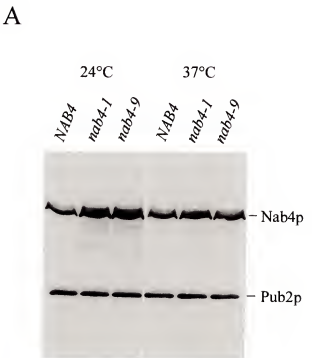
Figure 20. The overexpression of Nab4 protein in *nab4* mutants is due to the amino acid substitutions and not the copy number or episomal expression of mutant *nab4* alleles. Immunoblot anlaysis was performed on total protein isolated from cells grown in YPD at 24°C. The *nab4-1* allele was integrated at the *NAB4* chromosomal locus (*NAB4*, LDY133 v. YAO421), at the *TRP1* locus (*TRP1*, YAO405-2A v. YAO420), or expressed episomally from a single-copy plasmid (YCp50) or pRS316 (pRS; 2-5 copies/cell). The blot was probed with mAbs 3H1 and 2B1 against Nab4p or Pub2p, respectively. Sizes are indicated in kilodaltons.


NAB4
YCp50
pRS
TRP1
NAB4
nab4-1
NAB4
nab4-1
NAB4
nab4-1
NAB4
nab4-1
97
68
43
29
18
14
Nab4p
Pub2p

loss-of-function *nab4* mutants with which to begin directly investigating the role of Nab4p *in vivo*.

However, before continuing the functional analysis of Nab4p, it was important to address the elevated production of mutant Nab4 proteins. Although *nab4-1*, *nab4-4*, and *nab4-7* were each recessive to *NAB4*, the increased concentration of Nab4 protein in these cells could exacerbate or lead to defects in addition to those resulting from the loss of Nab4p function. I predicted, however, that any defects arising in this fashion would be constitutive, being observed during both permissive and non-permissive conditions. To try to circumvent this problem during functional studies, I attempted to generate a Ts^- *nab4* allele from which Nab4 protein was produced at wild type levels. I reasoned that the multiple missense mutations in the previously described *nab4* mutants might stabilize the mutant protein, perhaps by affecting the post-translational modifications to Nab4p. To eliminate or at least minimize this possibility, a mutant *nab4* allele was constructed that contained a single amino acid substitution, a leucine to serine at position 205 (Figure 18). This allele, designated *nab4-9* and first described as *hrp1-5* by Kessler et al. (1997), was chosen because it was reported to be Ts^- at 37°C and it would ultimately allow me to compare the results of my functional analyses on *NAB4* with those of others (Kessler et al., 1997). The *nab4-9* allele led to an inhibition of growth in culture with kinetics similar to the other *nab4* mutants described, and was recessive to *NAB4* (A.M. Krecic and M.S. Swanson, unpublished data). As observed with other *nab4* mutants, *nab4-9* overexpressed mutant protein at both permissive and non-permissive temperatures (Figure 21A). Therefore, there was a direct correlation between conditional growth and the level of Nab4p production in all *nab4* mutants tested.

Figure 21. The *nab4-9* allele is identical to other *nab4* alleles in protein expression and *ACT1* mRNA distribution. (A) Immunoblot analysis of total proteins form *nab4-9* (YAO467-1A) compared to wildtype (YAO417-1D) and *nab4-1* (YAO418-1D) grown in YPD at 24°C or 37°C for 2 h. (B) RNA blot analysis of *nab4-9*. Total RNA was isolated from *nab4-9* cells grown as in (A).


A
24°C
37°C
NAB4
nab4-1
nab4-9
NAB4
nab4-1
nab4-9
Nab4p
Pub2p
B
NAB4
nab4-9
0
2
0
2
1.6 kb
mRNA

When examined for the subcellular distribution of Nab4p, mutant Nab4 proteins, like the wild type protein, were predominantly nuclear at 24°C (Figure 22). When cells were shifted to 36°C, wild type Nab4 protein was still found in the nucleus, but about half of the protein accumulated in the cytoplasm in a pattern similar to that observed during mitochondrial induction (Figure 22). When *nab4* mutants were shifted to 36°C, only *nab4-1* exhibited a similar cytoplasmic distribution of Nab4 protein. In contrast, Nab4 protein expressed from *nab4-4* or *nab4-7* accumulated in the cytoplasm, but in a diffuse pattern. Since the total amount of Nab4 protein in these cells was already 2-3 times the normal level, the nuclear pool of Nab4p in mutant cells at 36°C was probably close to normal wild type levels. Each of the *nab4* mutants resulted in amino acid substitutions within the conserved RBDs of Nab4p (see Figure 18), suggesting that these regions of the protein may influence the subcellular localization of Nab4p. Changes in the primary structure of these domains may result in altered RNA-binding affinities or binding specificity of Nab4 protein. Nab4p is a nucleocytoplasmic shuttling protein and is likely to be a component of exported mRNP complexes that rapidly dissociates from the complex and is recycled to the nucleus. During specialized growth conditions when Nab4p accumulates in the cytoplasm of cells, Nab4p may remain associated with a subpopulation of mRNAs to facilitate cytoplasmic trafficking of these mRNAs or to help confer stability. Mutant Nab4 proteins may also be exported with mRNP complexes but, once in the cytoplasm, are unable to dissociate from the mRNP, resulting in a net loss of Nab4p from the nucleus. In this model, it would appear that the amino acid substitution of isoleucine to threonine at position 163 was important. This was the only mutation in common between *nab4-4* and *nab4-7*, which both showed altered re-distribution of Nab4

Figure 22. Mutant Nab4 proteins accumulate in the cytoplasm of cells at the non-permissive temperature. The localization of Nab4 mutant proteins was determined by indirect cellular immunofluorescence in cells grown at 24°C or shifted to 36°C for 2 h, using mAb 3H1. Shown at right are *NAB4* (YAO417-1D) and *nab4-1* (YAO418-1D).

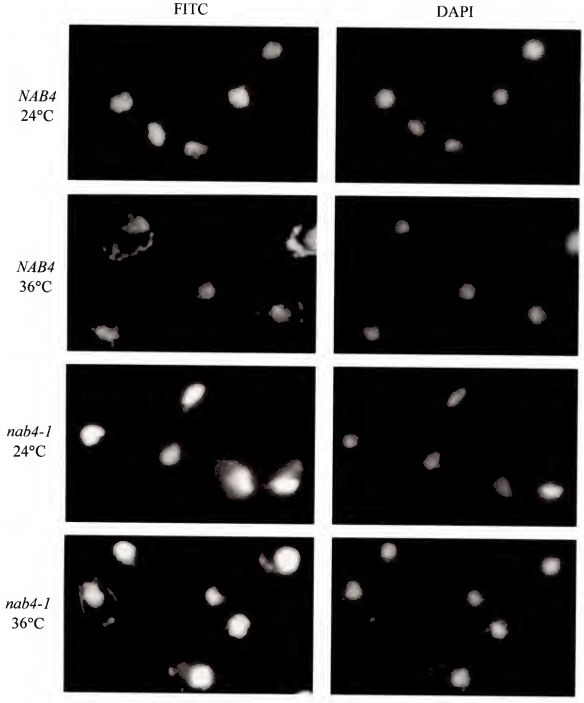
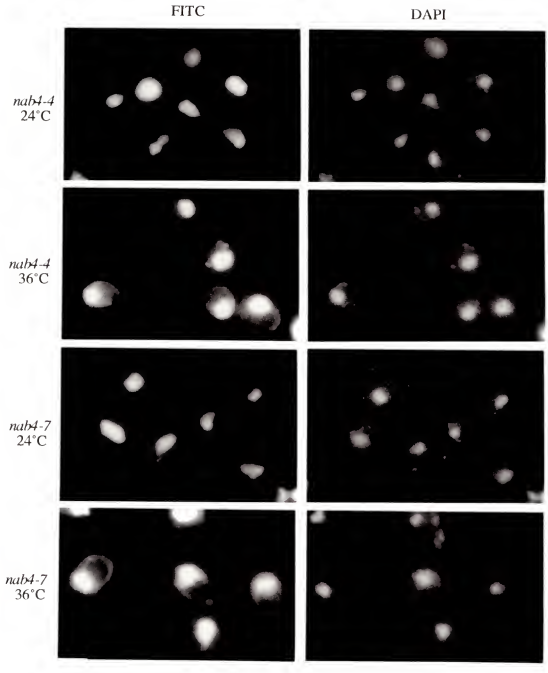

FITC
DAPI
NAB4
24°C
NAB4
36°C
nab4-1
24°C
nab4-1
36°C

Figure 22-continued. Mutant Nab4 proteins accumulate in the cytoplasm of cells at the non-permissive temperature. The localization of Nab4 mutant proteins were determined by indirect cellular immunofluorescence in cells grown at 24°C or shifted to 36°C for 2 h, using mAb 3H1. Shown at right are *nab4-4* (YAO428-1B) and *nab4-7* (YAO431-1C).


FITC
DAPI
nab1-4
24°C
nab1-4
36°C
nab1-7
24°C
nab1-7
36°C

protein, whereas *nab4-1* did not contain this mutation and exhibited normal re-distribution of Nab4 protein. Alternatively, mutant forms of Nab4p may simply be defective for nuclear localization at the non-permissive temperature. However, this is not likely, since the sequences necessary for nuclear localization of Nab4p were found in the carboxyl terminus, a region that is not mutated in any of these *nab4* alleles.

Mutations in *nab4* Do Not Cause a Block to mRNA Export or Pre-mRNA Splicing

I began the functional analysis of *nab4* mutants by determining whether Nab4p had a direct role in mRNA export. The distribution of poly(A)$^+$ RNA in cells was again visualized by hybridization with oligo(dT)$_{50}$ at 24°C and after shift to 36°C for 2 h (Figure 23). When compared to wild type, the mutant *nab4-1* strain exhibited no differences in poly(A)$^+$ RNA distribution at either temperature. Similar results were observed at various times after shift to 36°C ranging from 15 min to 8 h for *nab4-1*, *nab4-4* and *nab4-7*. This is in contrast to the results of overexpressing wild type Nab4p, indicating that, although Nab4p may be a component of exported mRNPs, it is not directly required for the export process.

Again, mRNA production was assessed by total RNA blot analysis. After a shift to the non-permissive temperature, *nab4* mutant strains showed the accumulation of longer transcripts for a variety of genes. This was most striking with *ACT1*, but was also observed for other genes including *CUP1* and *CYH2* (Figure 24; A.M. Krecic and M.S. Swanson, unpublished data). When compared to the *prp2-1* mutant for defects in pre-mRNA splicing, *nab4* mutants showed no significant accumulation of *ACT1* pre-mRNA. Although the levels of *CYH2* pre-mRNA seemed to increase relative to the amount of

Figure 23. Mutations in *NAB4* do not affect mRNA export. The distribution of poly(A)$^+$ RNA in cells was determined by in situ hybridization with oligo(dT)-digoxigenin. Wildtype (YAO417-1D) or *nab4-1* (YAO418-1D) cells were grown at 24°C or shifted to 36°C for 2 h. Poly(A)$^+$ RNA signals are compared to the DAPI-stained DNA.

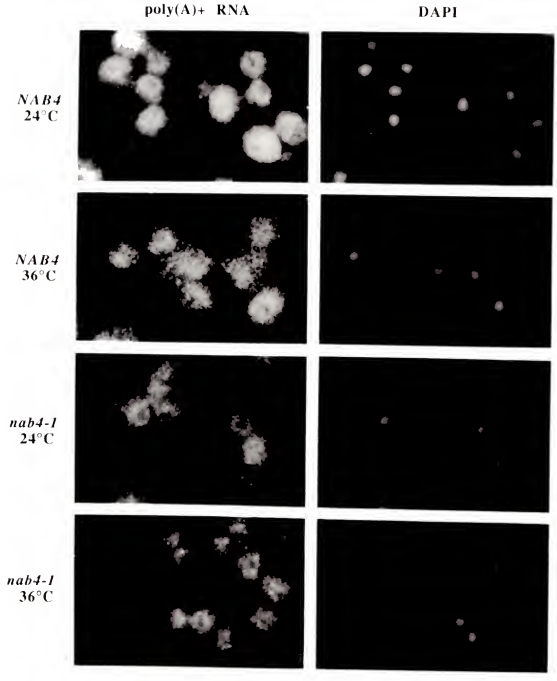
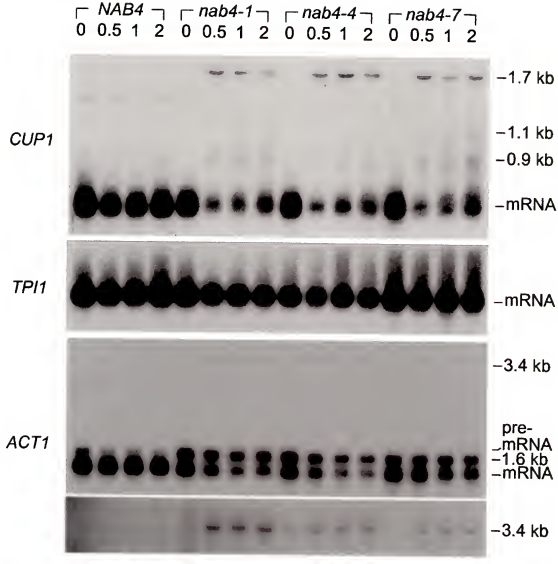

poly(A)+ RNA
DAPI
NAB4
24°C
NAB4
36°C
nab4-1
24°C
nab4-1
36°C

Figure 24. Mutations in *NAB4* lead to the accumulation of long transcripts for two genes. RNA blot analysis was performed on total RNA isolated from wildtype and *nab4* mutants grown at 24°C or shifted to 37°C for the indicated number of hours. Hybridization probes were used according to the Materials and Methods. The normal mRNAs, and the position of *ACT1* pre-mRNA, are indicated. The yeast strains used were those characterized in Figure 19.


NAB4
0 0.5 1 2
nab4-1
0 0.5 1 2
nab4-4
0 0.5 1 2
nab4-7
0 0.5 1 2
CUP1
–1.7 kb
–1.1 kb
–0.9 kb
–mRNA
TPI1
–mRNA
ACT1
–3.4 kb
pre-
–mRNA
–1.6 kb
–mRNA
–3.4 kb

CYH2 mRNA, this probably does not reflect a deficiency in splicing. Alternatively, the accumulation of longer transcripts, which I speculated were inappropriately produced or processed pre-mRNAs, suggests that the normal turnover pathways for such pre-mRNAs are only capable of eliminating a finite amount of pre-mRNA at any given time. Since the levels of *CYH2* pre-mRNA are regulated by a cytoplasmic nonsense-mediated degradation pathway, the higher steady-state level of *CYH2* pre-mRNA may be due to the inefficiency of an already saturated degradation pathway. Also, while mRNA levels initially decrease upon shift to 37°C, this occurs in wild type cells and is likely due to the general down-regulation of transcription during a transient heat-shock response. It was clear from these studies that unlike metazoan hnRNP A/B proteins, Nab4p was not involved in the regulation of splice site selection during pre-mRNA processing. There are relatively few intron-containing genes in *S. cerevisiae* and an apparent absence of any alternative splicing. Therefore, pre-mRNA splice site selection in yeast may not require regulatory factors beyond the basal splicing machinery. Hence, any conserved functional activity between Nab4p and metazoan hnRNP A/B proteins may be subtler than originally proposed.

Mutations in *nab4* Lead to Alterations in 3'-Endonucleolytic Cleavage During Pre-mRNA 3'-End Formation *In Vivo*

As shown in Figure 23, *ACT1* and *CUP1* show the accumulation of multiple longer transcripts, and with the exception of the 1.6 kb *ACT1* transcript, these appear within 30 min following shift to 37°C. The 1.6 kb *ACT1* transcript, barely detectable in wild type strains, was at least as abundant as the normal 1.4 kb *ACT1* mRNA in *nab4* mutants, including *nab4-9* (Figure 21B), at both 24°C and 37°C. Primer extension

analysis from *ACT1* transcripts revealed a second, but minor, initiation site approximately 160 nt upstream of the normal initiation site (Figure 25A). However, utilization of this site in the *nab4-1* strain was no more prevalent than in the wild type, representing about 20-25% in each. While this observation could account for the low-abundance 1.6 kb *ACT1* transcript found in *NAB4* cells, it cannot account for the abundance of this species found in *nab4* mutants. The 1.6 kb transcript was not a splicing variant, as it was not detected by hybridization using an *ACT1* intron-specific DNA probe (Figure 25B).

The accumulation of longer transcripts from both *ACT1* and *CUP1* has been observed previously in concert with aberrations in pre-mRNA 3'-end formation (Mandart and Parker, 1995; Forrester et al., 1992; Chanfreau et al., 1996). Therefore, it was hypothesized that the accumulation of longer transcripts in *nab4* mutants indicated a defect in pre-mRNA 3'-end processing. Previous work had shown that multiple *ACT1* 3'-ends are utilized for polyadenylation, and site usage is altered by mutations in genes required for 3'-end processing (Mandart and Parker, 1995). Initial reverse transcription-PCR (RT-PCR) studies confirmed the presence of longer *ACT1* 3'-UTRs in poly(A)$^+$ RNA isolated from *nab4* mutant strains grown at 24°C (A.M. Krecic and M.S. Swanson, unpublished data). This analysis was only semi-quantitative, but it appeared that the longer *ACT1* 3'-UTRs were more abundant in *nab4* mutants than in the wild type. To precisely map the 3'-ends of *ACT1* transcripts generated *in vivo*, a cDNA library was constructed using poly(A)$^+$ RNA purified from both *NAB4* and *nab4-1* strains grown at 24°C. The cDNA 3'-ends were amplified by PCR and clones corresponding to the 1.4 kb and 1.6 kb mRNAs were isolated and sequenced (Figure 26). The major polyadenylation site utilized in the *NAB4* strain was approximately 100 nt downstream of the translation

Figure 25. Defects in *ACT1* processing are not due to differential translation initiation or alternative splicing. (A) Primer extension analysis was performed on poly(A)$^+$ RNA from *NAB4* (YAO417-1D) and *nab4-1* (YAO418-1D) grown at 24°C. The sequence ladder was generated using pACT1.1 and numbers indicate nucleotides upstream of the *ACT1* initiating ATG. (B) RNA blot analysis of poly(A)$^+$ RNA from *nab4-1* and *prp2-1* cells grown at 24°C or shifted to 37°C for 2 h. Blots were hybridized with probes containing intron and exon sequences (I/E) or intron sequence only (I), as described in Materials and Methods.

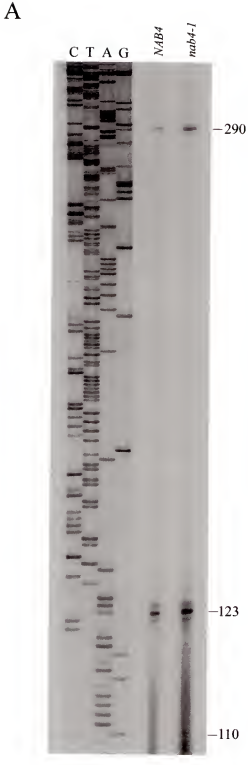
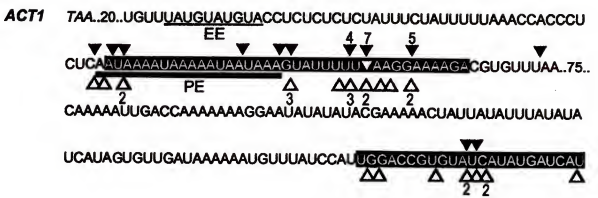

A
C T A G
NAB4
nab4-1
290
123
110
B
I/E
I
nab4-1
prp2-1
nab4-1
prp2-1
pre-mRNA
1.6 kb
mRNA

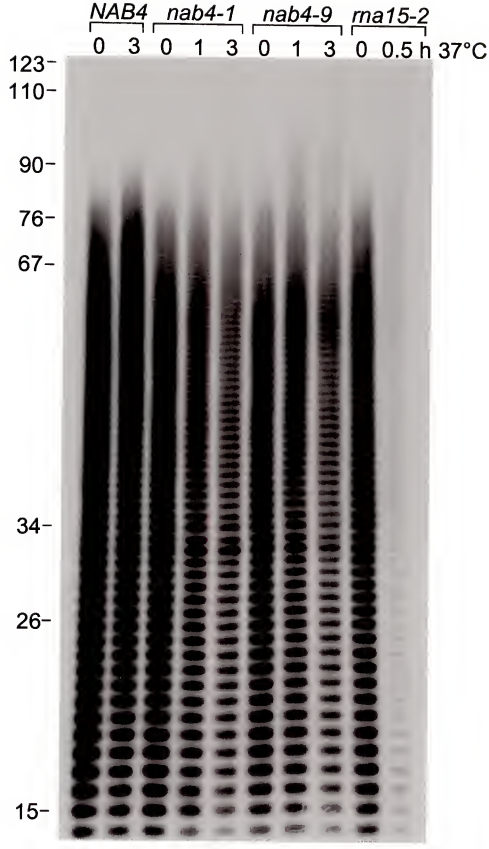
Figure 26. Mutations in *NAB4* lead to defects in the 3'-end processing of *ACT1*. The sequence shown is of the 3'-UTR of *ACT1* and the stop codon is italicized. The efficiency (EE) and positioning (PE) elements are indicated by light and heavy underlines, respectively. The 3' ends of *ACT1* were mapped *in vivo* from poly(A)$^+$ RNA isolated from cells grown in YPD at 24°C, as described in the Materials and Methods. Closed triangles represent the 3' ends of clones isolated from *NAB4* cells (YAO417-1D) and open triangles represent those clones recovered from *nab4-1* cells (YAO418-1D). Numerals above and below triangles indicate the number of clones isolated. Clones represented by triangles with no numeral were isolated once.


ACT1
TAA..20..UGUUUAUGUAUGUACCUCUCUCUCUAUUUCUAUUUUUAAACCACCCU
EE
4 7
5
CUCAAUAAAAUAAAAUAAUAAAGUAUUUUUAAGGAAAAGACGUGUUUAA..75..
PE
2
3
3 2
2
CAAAAAUUGACCAAAAAAAGGAAUAUAUAUACGAAAAACUAUUAUAUUUAUAUA
UCAUAGUGUUGAUAAAAAAUGUUUAUCCAUUGGACCGUGUAUCAUAUGAUCAU
2 2

termination codon, consistent with the position of the major site previously reported. In contrast, and as predicted from the RNA blot analysis and RT-PCR analysis, the *nab4-1* mutant showed significant cleavage at sites in a region approximately 300 nt downstream of the termination codon. This difference in the position of cleavage sites could account for the difference between the 1.4 kb and 1.6 kb *ACT1* mRNAs found in all *nab4* mutants and confirmed that the selection of 3'-end endonucleolytic cleavage sites is altered constitutively by mutations in *NAB4*. These downstream sites are also consistent with those reported by Mandart and Parker (1995). However, the utilization of these sites in *nab4* mutants appears to be distributed in a broader cluster around both the proximal and distal sites previously identified. This observation emphasizes the loss of specificity in cleavage site selection exhibited by *nab4* mutants.

As stated above, the utilization of alternative downstream cleavage sites has been reported for a variety of 3'-end processing mutants, including *pap1-1*, *rna14-3 and rna15-2* (Mandart and Parker, 1995). These mutants also show rapid loss of mRNAs and decreases in poly(A) tail length distribution when grown at the non-permissive temperature for these strains (Minvielle-Sebastia et al., 1991). However, when tested for polyadenylation defects *in vivo*, there was no significant effect on poly(A) tail lengths in any of the *nab4* mutants up to 3 h after shift to the non-permissive temperature (Figure 27). This was very different from *rna15-2*, which showed almost complete loss of poly(A) tails within 30 min. Consistent with the decreased growth rate compared to wild type at 37°C, all *nab4* mutants showed a global decrease in the level of all tail lengths, but no specific decrease in the average maximal length of poly(A) tails. In contrast to recent reports that Nab4p/Hrp1p is required for both cleavage and polyadenylation (see

Figure 27. Mutations in *NAB4* do not exhibit a poly(A) tail length defect. Poly(A) tail analysis was performed using RNAs isolated from cells grown at 24°C or after shift to 37°C for the indicated number of hours (h). The *NAB4* strains used were those characterized in Figure 19. The *rna15-2* strain (LM45) was kindly provided by Lionel Minvielle-Sebastia. Markers are MspI-cut pBR322 and are indicated in nucleotides.


NAB4
nab4-1
nab4-9
rna15-2
0 3 0 1 3 0 1 3 0 0.5 h 37°C
123
110
90
76
67
34
26
15

below), it appeared here that Nab4p is required for correct 3'-end formation *in vivo* but not required at the level of polyadenylation. Since *nab4* mutants appeared to affect 3'-end cleavage of transcripts from a subset of genes, it was possible that *nab4* mutants did affect polyadenylation, but only for some transcripts. A transcript-specific effect such as this could not be ruled out by the analysis of poly(A) tails from total cellular RNA shown here. However, it was clear that the defect in cleavage specificity was constitutive in *nab4* mutants, indicating that the role of Nab4p in 3'-end processing was quite different from others factors involved in cleavage and polyadenylation.

Nab4p Regulates 3'-Endonucleolytic Cleavage Site Selection During Pre-mRNA 3'-End Formation

During the course of these studies, Nab4p/Hrp1p was identified as the sole component of CF IB, a biochemically purified factor reportedly required for *in vitro* cleavage reactions as well as for polyadenylation both *in vitro* and *in vivo* (Kessler et al., 1996; Kessler et al., 1997). This finding was in disagreement with the observations made using *nab4* mutants, including *nab4-9*, which is identical to *hrp1-5* (Kessler et al., 1997). These *nab4* mutants displayed only minor temperature-dependent defects in 3'-end cleavage *in vivo* and no defects on polyadenylation. To test the requirement of Nab4p in 3'-end processing further, extracts were prepared from *NAB4* and *nab4* mutant strains and assayed for the ability to cleave or polyadenylate a *CYC1* precursor RNA *in vitro* (R.E. Hector and M.S. Swanson, unpublished data). All of the mutant extracts exhibited near wild type levels of cleavage activity at a temperature permissive for growth of these mutants, and decreases in activity at the non-permissive temperature were 50% or less (Table 2). Likewise, *nab4* mutant extracts were also competent for polyadenylation of

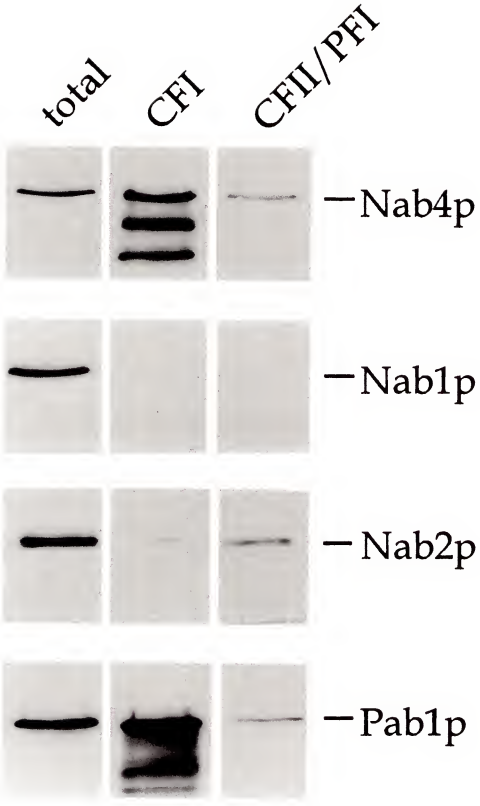
the *CYC1* 5' cleavage fragment, generated first by pre-cleaving the full-length *CYC1* precursor with a wild type extract. Polyadenylation was less efficient in both *nab4-1* and *nab4-7* extracts, but this was true at both permissive and non-permissive temperatures, and there was no additional temperature-dependent loss of activity at the elevated temperature when compared to the wild type extract. Immunoblot analysis of these extracts suggested that the activity of other proteins that influence polyadenylation, such as Pab1p and Nab2p, might be functionally compromised by proteolysis during preparation of the extracts, and thereby decrease the efficiency of polyadenylation (R.E. Hector and M.S. Swanson, unpublished data).

Table 2. *In vitro* cleavage of *CYC1* by *NAB4* and *nab4* extracts. Relative activity of extracts is expressed as percent of wild type at 24°.

	24°C	36°C
NAB4	100	148
nab4-1	113	62
nab4-4	135	89
nab4-7	74	59

That Nab4p was an abundant component of CF I was confirmed by immunoblot analysis using a monoclonal antibody specific for Nab4p (Figure 28). This was in contrast to low levels of Nab4p in a crude preparation of CF II that also contained PF I. The purification of Nab4p with CF I was also specific for Nab4p, as Nab1p/Npl3p and Nab2p were not present in CF I. To begin to reconcile the differences between these data and those of Kessler et al., a collaboration was initiated with Lionel Minvielle-Sebastia

Figure 28. Nab4p is a component of yeast CF I. Immunoblot analysis of CF I and CF II, probed with mAbs against the indicated proteins as described in Materials and Methods. CF I and partially purified CF II (CF I/PFI) were a gift from Claire L. Moore and Marco M. Kessler. Approximately 1.2 μg of CF I and 1.5 μg of CF II were analyzed. Total protein extract was prepared from YJA501 grown to mid-log phase in YPD at 30°C.

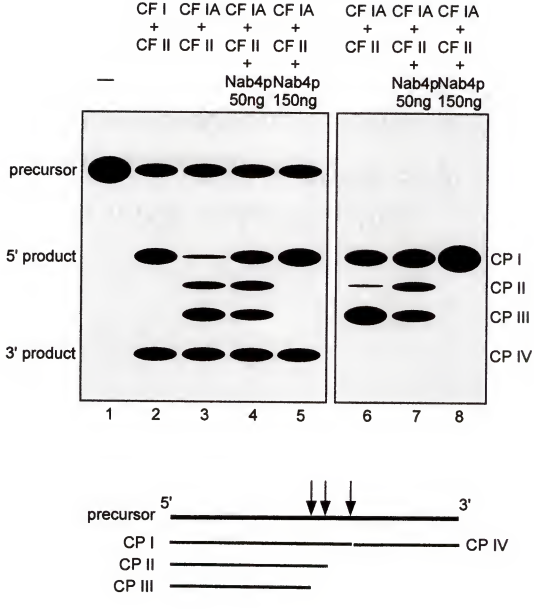

total
CFI
CFII/PFI
Nab4p
Nab1p
Nab2p
Pab1p

and Walter Keller of the University of Basel, Switzerland. Keller and co-workers are experts in the field of 3'-end formation and have the expertise to address the role of Nab4p in 3'-end formation by using purified cleavage and polyadenylation factors.

The results of Kessler et al. were produced using CF IA, CF II and recombinant Hrp1p. The CF IA and CF II fractions used in our studies were purified and assayed for the presence of Nab4p. Immunoblot analysis using a monoclonal antibody revealed no detectable Nab4p in either the CF IA or CF II fractions in contrast to CF I and PF I (L. Minvielle-Sebastia, personal communication). The cleavage activity of CF IA and CF II was tested on a *CYC1* 3'-end processing precursor (Figure 29, lane 1) and the results are summarized in Figure 29 (L. Minvielle-Sebastia, personal communication). In agreement with our previous results using crude cell extracts from *nab4* mutants, CF IA and CF II, devoid of Nab4p, were sufficient for cleavage of *CYC1* (Figure 29, lane 3). Surprisingly, additional cleavage products were also generated from *CYC1* in the absence of Nab4p in contrast to normal cleavage with CF I and CF II (Figure 29, compare lanes 2 and 3). Adding recombinant Nab4p to these reactions appeared to prevent the formation of these additional cleavage products and restored normal cleavage (Figure 29, lanes 4 and 5). Similar similar results were obtained with other 3'-end processing precursors, including GAL7 and the yeast Ty element U3RU5 polyadenylation site (L. Minvielle-Sebastia, personal communication). From these observations, we concluded that Nab4p was not required for 3'-end endonucleolytic cleavage but could influence the fidelity of cleavage.

In the case of *CYC1*, three major cleavage products were generated in the absence of Nab4p (Figure 29, lane 3). These included what appeared to be the normal 3' cleavage product, designated cleavage product IV (CP IV), the most abundant fragment, CP III,

Figure 29. Summary of *in vitro* cleavage of *CYC1* precursors in the presence and absence of Nab4p. Shown are the results of *in vitro* cleavage reactions using ^{32}P-labeled *CYC1* 3'-end formation substrates and various combinations of purified cleavage factors CF I and CF II, and GST-Nab4p, as indicated above individual lanes. Reaction products were analyzed by separation on denaturing polyacrylamide/urea gels followed by autoradiography. A representative autoradiograph is schematized, indicating the normal 5' and 3' cleavage products (also referred to as CP I and CP IV, respectively) as well as additional cleavage products termed CP II and CP III. The line drawing below illustrates the full-length *CYC1* precursor RNA and the various cleavage products observed in these studies. Arrows above the precursor indicate the sites at which cleavage is proposed to occur to give rise to the various cleavage products.


CF I + CF II
CF IA + CF II
CF IA + CF II + Nab4p 50ng
CF IA + CF II + Nab4p 150ng
CF IA + CF II
CF IA + CF II + Nab4p 50ng
CF IA + CF II + Nab4p 150ng
—
precursor
5' product
3' product
CP I
CP II
CP III
CP IV
1 2 3 4 5 6 7 8
precursor
5'
3'
CP I
CP IV
CP II
CP III

which is ~35 nt shorter than the normal 5' fragment, CP I, and CP II. The addition of Nab4p eliminated the production of CP II and CP III in a concentration-dependent manner (Figure 29, compare lanes 4 and 5 to lane 3). At the highest concentrations of Nab4p (Figure 29, lane 5), the formation of only CP I and CP IV was observed. As increasing amounts of Nab4p were titrated into the cleavage reaction, there was a progressive increase in the amount of CP I concomitant with the loss of CP II and CP III. At the same time, the amount of CP IV showed only a modest change. These results suggested that CP II and CP III may be produced by subsequent cleavage of CP I. To test this idea, pre-cleaved CP I was isolated from a complete cleavage reaction and then used as substrate for cleavage by CF IA and CF II. In the absence of Nab4p, CP III appeared at high levels along with a very low amount of CP II (Figure 29, lane 6). This indicated that the additional cleavage fragments are generated from CP I, presumably by the utilization of cryptic cleavage sites within CP I, upstream of the normal polyadenylation site. Again, the subsequent cleavage of CP I was inhibited in response to increasing concentrations of Nab4p (Figure 29, lanes 7 and 8).

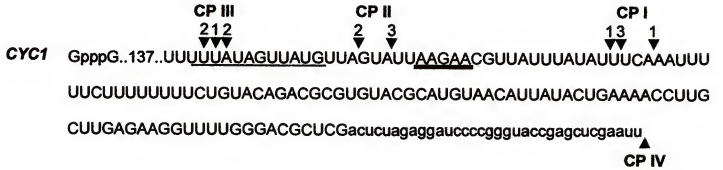
The cleavage of precursors derived from *GAL7* and U3RU5 was also examined in the absence of Nab4p and with increasing concentrations of Nab4p. Because the normal *in vitro* cleavage patterns of *GAL7* and U3RU5 are more complex than that of *CYC1*, the results of these experiments were more difficult to interpret, but were nevertheless, consistent with a role for Nab4p in ensuring the fidelity of cleavage site selection during 3'-end formation.

The relationship between cleavage and Nab4p became clearer by examining the time-course of *CYC1* (L. Minvielle-Sebastia, personal communication). Cleavage

reactions were performed for 0-90 min using purified CF IA and CF II devoid of Nab4p, revealing a product-precursor relationship among CP I, CP II and CP III. CP I appeared quite rapidly, within 0.5 min. The levels of CP I remained steady out to 12-14 min as CP II and CP III began to accumulate. By 60 min, there was a significant loss of CP I, while CP II was maximal at 6-7 min and CP III continued to accumulate throughout the time-course, as did CP IV, the normal 3' cleavage fragment. These results suggested that cleavage initially occurred to produce the normal 5' cleavage fragment, CP I. Subsequent cleavage of CP I then resulted in the sequential production of CP II and CP III. At the same time, the accumulation of CP IV was unaffected.

To gain insight into the mechanism of cleavage, the 3' ends of CP I-IV fragments of *CYC1* were determined (Figure 30). RNA cleavage fragments were isolated and used to produce cDNAs. Cloning and sequencing of these cDNAs confirmed the prediction, that the 3' ends of CP II and CP III were within 35 nt upstream of the normal (CP I) cleavage site. The additional cleavage sites were specific and, as implied from the previous time-course studies, were utilized in a 3'→5' direction. Since CP I was sequentially cleaved to produce CP II and CP III, it appeared that the cleavage mechanism was processive rather than distributive. The mechanism of cleavage also appeared to be unidirectional, since the 3' cleavage fragment (CP IV) was not further cleaved. Alternatively, CP II and CP III may be degradation products of CP I resulting from an exonuclease activity. However, this is unlikely due to the specific accumulation of CP II and CP III rather than a broader population of less abundant cleavage products ranging in size between CP I and CP III or shorter. The possibility of an exonuclease activity was eliminated by separating cleavage reactions devoid of Nab4p on longer gels,

Figure 30. Mapping of *CYC1* cleavage products generated *in vitro* in the absence of Nab4p. The sequence of the *CYC1* precursor RNA is shown with lowercase letters representing pG4-CYC1 plasmid sequence resulting from run-off transcription. The efficiency (EE) and positioning (PE) elements are indicated by light and heavy underlines, respectively. The 3' ends of *CYC1* cDNA clones representing CP I-CP III isolated as described in the Materials and Methods, are indicated by closed triangles. Numerals above triangles indicate the number of clones isolated. The 3'end of CP IV, generated by run-off transcription, is also indicated.


CP III
CP II
CP I
2 1 2
2 3
1 3 1
CYC1
GpppG..137..UUUUUUAUAGUUAUGUUAGUAUUAAGAACGUUAUUUAUAUUUCAAAUUU
UUCUUUUUUUUCUGUACAGACGCGUGUACGCAUGUAACAUUAUACUGAAAACCUUG
CUUGAGAAGGUUUUGGGACGCUCGacucuagaggaucccgggguaccgagcucgaauu
CP IV

allowing the resolution of the smaller fragments produced by cleavage of CP I to CP II (15 nt) and CP II to CP III (20 nt). Therefore, cleavage appeared to occur by means of a processive 3'→5' endonuclease, the fidelity of which was regulated by Nab4p. Still, it is possible that the additional cleavage fragments are formed by a rapid 3'-5' exonuclease which pauses at these specific upstream sites.

Thus far, these studies have allowed me to present the working model that Nab4p is involved in modulating the utilization of normal and cryptic endonucleolytic sites during pre-mRNA 3'-end formation. This activity was first suggested by observations made *in vivo* and strongly supported by extensive *in vitro* studies. How, and the extent to which, Nab4p might contribute to gene expression, specifically during the 3'-end formation of pre-mRNAs, remain to be explored. Numerous attempts were made to establish a system by which to investigate the effect of depleting Nab4p *in vivo*. However, because of the toxicity of relatively small increases in Nab4p concentration, this proved to be difficult. One possible alternative approach to depleting Nab4p would be to generate a conditional *NAB4* allele by constructing a gene fusion between Nab4p and the thermo-labile dihydrofolate reductase, which is rapidly degraded at the non-permissive temperature (Dohmen et al., 1994). Fusion proteins produced in this way are also highly temperature-labile and may provide a means of regulating Nab4p production. The effect of Nab4p depletion could then be examined by shifting cells to the non-permissive temperature and avoid the caveats encountered with *nab4* Ts^- mutants in these studies. These experiments have not been performed but will be important as an independent approach to evaluating the function of Nab4p *in vivo*.

DISCUSSION

I have described here the isolation and characterization of Nab4p, a yeast hnRNP both structurally and functionally related to the most abundant class of human hnRNPs, the hnRNP A/B proteins. The existence of such a protein in yeast suggests that hnRNP A/B-type proteins possess an essential function, probably in mRNA biogenesis, that has been conserved throughout evolution. The identification of an hnRNP A/B-type protein in yeast has provided a powerful new genetic system to help elucidate the functions of this major class of hnRNPs.

My research has revealed a number of unique features of Nab4p when compared to previously isolated yeast hnRNPs. Like other yeast hnRNPs, Nab4p is essential for cell viability. However, the intracellular concentration of Nab4p must be carefully controlled to maintain normal cellular physiology. Nab4p is a concentration-dependent regulator of 3'-end processing, the first such factor identified in yeast, and is the first hnRNP to be ascribed a direct role in 3'-end processing. Finally, the specific cytoplasmic accumulation of this hnRNP, and not other yeast shuttling hnRNPs, provides strong support for the existence of novel cytoplasmic functions for hnRNPs.

Functional Clues Based on the Importance of the Intracellular Concentration of Nab4p

That the intracellular concentration of Nab4p must be so carefully controlled was quite unexpected given the results of overexpressing other Nab proteins. The only

growth inhibition due to overexpressing other Nab proteins was with Npl3p/Nab1p, and the inhibitory effect was mild compared to that observed with Nab4p overproduction (Flach et al., 1994). These observations first suggested that the function of Nab4p might be dependent upon the concentration of Nab4p relative to other components involved in the same process. This is the case for hnRNP A/B proteins in the selection of splice sites during pre-mRNA splicing (Mayeda and Krainer, 1992; Cáceres et al., 1994). If the function of Nab4p is dependent upon interaction with another factor, then altering the level of Nab4p may change the stoichiometry of the interaction and alter the activity or specificity of Nab4p.

Alternatively, overexpression of Nab4p may lead to a dominant-negative effect, in which higher levels of Nab4p result in a gain-of-function that disrupts a metabolic pathway related, or unrelated, to the function of Nab4p. There have been a number of studies specifically designed to identify genes whose *GAL1,10*-inducible overexpression causes lethality in yeast (Liu et al., 1992; Espinet et al, 1995; Akada et al., 1997). Each of these studies identified a different set of genes, which may be due to differences in the genetic background among the yeast strains employed. Nevertheless, these results have been confirmed independently for a number of genes, including *ACT1* (Liu et al., 1992) and *RBP1*, a glucose-repressible gene encoding a RNA-binding protein (Lee and Moss, 1993). In addition to these, many of the genes identified encode structural or morphogenesis-related proteins, multifunctional transcription factors, or proteins involved in signal transduction pathways. Others seemed to affect cell cycle regulation, leading to elongated-bud or growth-arrest phenotypes. These observations present a wide range of possible functions for proteins whose overexpression is toxic to cells.

Increasing the concentration of Nab4p may lead to Nab4p interactions with RNA or DNA sequences that would not occur under normal physiological conditions. For example, Nab4p contains more RNA-binding motifs than other Nab proteins, including the two highly conserved RBDs and two RGG box motifs, each of which has the ability to bind RNA. While Nab4p exhibits sequence-preference RNA binding as well as single-stranded DNA binding *in vitro,* this multiplicity of RNA-binding motifs may result in a larger variety of RNA or DNA sequences to which Nab4p can bind with a broad range of binding affinities. Lower affinity interactions may occur only when levels of Nab4p are elevated.

One possible *in vivo* target for such an interaction is telomeric sequences. Several proteins have been identified which bind to telomeric sequences either *in vitro* or *in vivo.* Among these are Rap1p, a transcription factor that is also involved in transcriptional silencing at telomeres (Longtine et al., 1989; Liu et al., 1994), and the single-stranded telomeric DNA repeat $(TG_{1\text{-}3})_n$-binding proteins Gbp2p and Nsr1p (Lin and Zakian, 1994). Of these, Rap1p is the only one that has an effect on telomere length or transcriptional silencing by telomere position effect. Overexpression of Rap1p is toxic and causes changes in the length of telomeric sequences (Chambers, 1996). Overproduction of Nsr1p, a NLS-binding protein nucleolar protein involved in ribosome biogenesis, is also toxic. The nucleolar localization of Nsr1p appears to be a function of its RNA-binding properties and not the result of a specific nucleolar targeting signal (Yan and Melese, 1993). Therefore, as increased levels of Nsr1p accumulate in the nucleus of cells and the nucleolar binding sites become saturated, Nsr1p may begin to bind elsewhere in the nucleus, perhaps at telomeres. Although elevated levels of Nsr1p would

appear to have no effect on telomere length or telomere position effect (Lin and Zakian, 1994), it may affect telomere function in a way not measured by these assays. Curiously, like Nab4p, Nsr1p binds RNA with striking similarity to hnRNP A1 (Gamberi et al., 1994), suggesting that at high enough concentrations, Nab4p may also have the capacity to bind telomeric repeats. This could explain distortions in the shape of the nucleus in cells observed during the overexpression of Nab4p, since telomeres are important in the nucleation of repressed chromatin near telomeres. Perturbations in telomere function or structure may result in localized disruption of chromatin structure, which could alter the overall nuclear structure. If a role related to telomeres were identified for Gbp2p, it would be important to explore the functional significance of the two-hybrid interaction between *NAB4* and *GBP2*. This analysis should determine whether Nab4p binds telomeric sequences *in vitro*. The sequence preference of Nab4p for poly(U) and poly(G) and the binding similarities among Nab4p, hnRNP A/B proteins and Nsr1p suggest that this is possible. Assays could then be performed using *nab4* mutants, as well as during overexpression of Nab4p, to look for specific effects on telomere length or telomere silencing that may be epigenetic to other observed defects.

Finally, overexpression of Nab4p may cause cell death indirectly, by sequestering and interfering with the function of a second protein. The overexpression of Nab4p caused two obvious phenotypes, the nuclear accumulation of poly(A)$^+$ RNA and the accumulation of aberrantly long transcripts. These observations suggest potential roles for Nab4p in pre-mRNA processing and mRNA export. The significance of these observations and possible mechanisms by which the overexpression of Nab4p may lead to these defects are discussed below.

Nab4p Is a Concentration-Dependent Regulator of 3'-End Cleavage Site Selection

I have reported here that Nab4p has a direct effect on selection of 3'-end cleavage sites *in vivo* using *nab4* mutants (Figure 26). This has been corroborated by reconstitution of normal cleavage *in vitro* (Figure 29). This effect was concentration-dependent since even modest increases in Nab4p alter cleavage site utilization. This feature of the role of Nab4p in the formation of 3'-ends may explain, in part, the precision with which Nab4p levels must be maintained. This feature might also account for the constitutive defect in *ACT1* 3'-end processing exhibited by *nab4* mutants that expressed 2-3 times the normal level of Nab4 protein. However, not all transcripts were affected by *nab4* mutants and there was only a modest temperature-dependent defect. The differences in sensitivities of different transcripts to the concentration of Nab4p may be a function of the relative strength of poly(A) signals or the presence of multiple cleavage sites within a pre-mRNA. Since Nab4p is not required for 3'-end endonucleolytic cleavage (Figure 29), these observations strongly suggest that regulating the specificity of endonucleolytic cleavage at the 3' end is an important function of Nab4p.

If formation of appropriate 3' ends is important to the regulation of gene expression, then alterations in cleavage site selection or the inability to utilize specific sites when necessary may have serious consequences for cellular metabolism. Recently, it was reported that changes in the abundance of the 64 kDa subunit of CstF (CstF-64) may regulate alternative polyadenylation during B lymphocyte activation (Takagaki et al., 1996). This and earlier precedents for such a phenomenon support the idea that the selection of endonucleolytic cleavage sites during 3'-end formation may be dose-

dependent on specific 3'-end processing factors (Mann et al., 1993, O'Hare 1995). Nab4p is the first factor of this sort identified in yeast, and the ability of Nab4p to regulate the utilization of cleavage sites may represent a conserved function between this protein and metazoan hnRNP A/B proteins, namely in the specification of endonucleolytic cleavage sites during pre-mRNA processing. It remains to be seen if Nab4p functions in this respect along with antagonists, as is the case with hnRNP A/B proteins and the SR proteins of mammalian cells during alternative pre-mRNA splicing, or if this activity is dependent on the relative ratio of Nab4p to other cleavage factors.

Increasing the concentration of Nab4p or mutation of *NAB4* had a surprising effect on 3'-end formation. Moreover, the overexpression of Nab4p appeared to exacerbate the defect on 3'-end formation, perhaps by sequestering CF IA or other 3'-end processing factors (see below), thus altering cleavage of all transcripts examined. The accumulation of these improperly processed transcripts may in turn represent the poly(A)$^+$ RNA that accumulates in the nucleus of cells during Nab4p overexpression.

This is the first case of an hnRNP influencing pre-mRNA 3'-end formation in eukaryotic cells. As noted previously, the mammalian hnRNP C protein preferentially binds to substrate RNAs containing 3'-end processing signals, but no direct functional consequence of this interaction has been reported. The lack of a consensus poly(A) signal in yeast draws into question how endonucleolytic cleavage is specified. The function of Nab4p as an auxiliary specificity factor may have evolved to overcome the lack of a consensus poly(A) site in yeast, by stabilizing the interaction of CF IA with favored cleavage sites. As a primary RNA-binding protein of nascent pol II transcripts, Nab4p may be the first 3'-end processing factor to associate with the 3' untranslated

region of a pre-mRNA. As such, Nab4p may recruit the basal cleavage factors to the pre-mRNA. Two-hybrid and synthetic-lethal interactions between Nab4p and both Rna14p and Rna15p have been reported (Kessler et al., 1997), suggesting a possible mechanism by which Nab4p could recruit CF IA. In addition, it appears that CF IA may represent a functional soluble complex that exists *in vivo*, since Nab4p is separable from CF I. Specific and stable interaction at the appropriate RNA elements may be a combination of Nab4p binding and Rna15p-mediated CF IA binding. There is no evidence that Nab4p or any component of CF IA interacts with CF II, but this interaction may not be necessary for recruitment of CF II to the cleavage complex since Cft2p appears to recognize the efficiency element. Although there are no known RNA-binding motifs in Cft2p, this 105 kDa subunit of CF II has homology to the 100 kDa subunit of mammalian CPSF and was the only component of CF II crosslinked to a full-length RNA substrate (Zhao et al., 1997). Crosslinking was dependent upon both the $(UA)_6$ efficiency element as well as sequences downstream of the poly(A) site. Nevertheless, it was clear from our results that cleavage requires only CF IA and CF II. It is clear from the mammalian system that CPSF, CstF and CF I_m all have the ability to bind polyadenylation precursors, but the most stable RNA-protein complex is one that contains all three of these factors. The same is probably true in yeast. However, the specific formation of this complex on mammalian pre-mRNAs seems to be largely determined by the almost invariant AAUAAA poly(A) signal. Therefore, an auxiliary specificity factor like Nab4p may not be necessary for correct utilization of cleavage sites in mammalian cells. An exception to this might be the requirement for tissue-specific or development-specific factors to

regulate alternative polyadenylation, akin to the effect that the abundance of CstF-64 has on alternative polyadenylation in B lymphocyte activation (Takagaki et al., 1996).

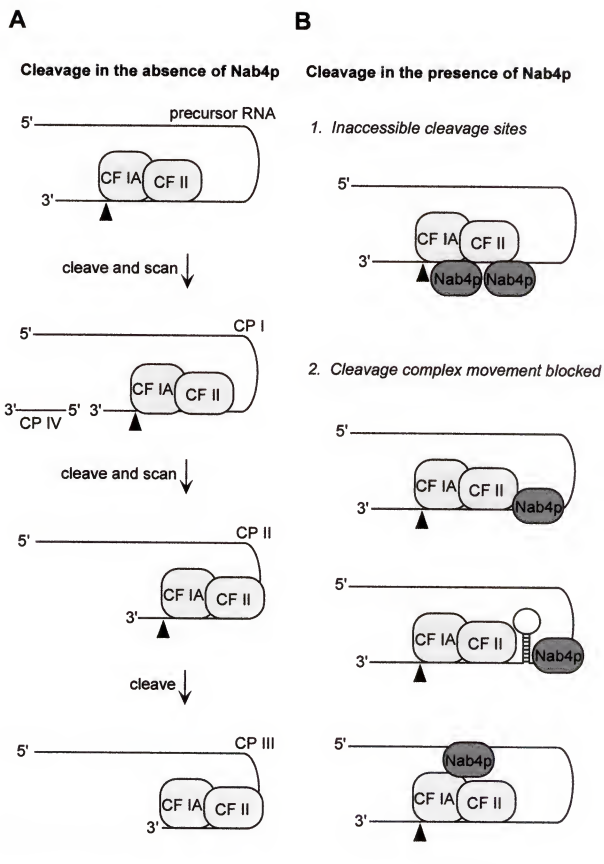
The observation that Hrp1p/Nab4p can be crosslinked to a precleaved substrate and that this activity is stimulated by CF IA suggest that Nab4p may serve a dual auxiliary role. Following cleavage, CF II may dissociate, since it is dispensable for polyadenylation. The Nab4p-CF IA complex bound to RNA remains in stable association with the substrate to aid in the recruitment of PF I and Pap1p (via interactions between Rna14p, Rna15p, Pcf11p and Fip1p, as described previously) for subsequent polyadenylation. If Nab4p is important for polyadenylation, the ability to separate CF I into CF IA and CF IB may also explain why CF I is required for both cleavage and polyadenylation. Yet the two steps are not as tightly coupled as in mammalian 3'-end processing. In this way, Nab4p may increase the efficiency with which polyadenylation is initiated, but may not be required for poly(A) addition.

These studies have revealed new aspects of the mechanism of cleavage during 3'-end formation in yeast. It has been widely held that this event occurs via a distributive reaction. However, the additional cleavage products and the precursor-product relationship observed among them during the time course of *CYC1* cleavage in the absence of Nab4p indicate that cleavage can occur in a processive manner. The discrete upstream cleavage fragments and the corresponding short, but detectable, downstream cleavage fragments further show that this processivity is the result of site-specific cleavage by an endonuclease and not an exonuclease. Although the source of this endonuclease activity remains unknown, these newly described features of the

endonucleolytic reaction will be important considerations when designing experiments to identify the endonuclease.

The processive nature of the 3'-end endonuclease in yeast indicates that the cleavage complex has the ability to move along the pre-mRNA substrate and, in the absence of Nab4p, this occurs in a 3'→5' direction. This is perhaps analogous to the scanning process that occurs prior to translation initiation by ribosomes during protein synthesis. Perhaps the cleavage complex enters the RNA downstream of the normal cleavage site and proceeds upstream along the RNA until it comes upon a cleavage site (Figure 31). The utilization of a given cleavage site, however, would depend upon the establishment of stable RNA-protein and protein-protein contacts within the context of the cleavage apparatus. In the absence of Nab4p, cleavage is more permissive, suggesting that Nab4p blocks certain sites from cleavage. Alternatively, Nab4p may be needed to stabilize the complex at a specific site to facilitate cleavage and to prevent subsequent cleavage of the upstream fragment. In the absence of Nab4p, the cleavage complex can remain associated with the upstream fragment, continuing to cleave at permissive sites. Eventually, the positioning and efficiency elements are cleaved away from the precursor and presumably the basal cleavage complex can no longer maintain stable associations with the RNA and cleavage ceases. This model predicts the presence of upstream cryptic sequence elements in *CYC1*, since cleavage can occur at positions upstream of the natural poly(A) site. The *cyc1-512* mutant (Zaret and Sherman, 1982) is deleted for 38 bp upstream of the normal cleavage site, CP I, which includes the apparent efficiency element for CP I. Hence, cleavage of this precursor is inefficient both *in vivo* and *in vitro* using crude polyadenylation extracts or purified CF I and CF II. The

Figure 31. Possible mechanisms for the role of Nab4p in the regulation of 3'-end endonucleolytic cleavage site selection during pre-mRNA 3'-end processing. (A) Cleavage of a precursor RNA in the absence of Nab4p. The cleavage complex assembles on the RNA and cleaves at a competent site (indicated by filled triangles). In the absence of Nab4p, the cleavage complex moves and subsequently cleaves at additional sites. (B) Cleavage of a precursor RNA in the presence of Nab4p. Nab4p may inhibit cleavage at additional sites by (1) binding to or near cleavage sites or (2) by inhibiting the movement of the cleavage complex. By binding to or near other cleavage sites, Nab4p may render these additional sites inaccessible to the cleavage complex or compete with CFIA or CF II for RNA binding, preventing them from binding to the precursor in a manner conducive to cleavage at a particular site. Alternatively, movement of the cleavage complex may be inhibited by Nab4 protein acting as a physical block on the RNA, by facilitating the formation of RNA secondary structures, or by interacting with the cleavage complex to stabilize its binding to the precursor, thus enhancing cleavage at a specific site.


A
Cleavage in the absence of Nab4p
precursor RNA
5'
3'
CF IA
CF II
cleave and scan
CP I
CP IV
CP II
cleave
CP III
B
Cleavage in the presence of Nab4p
1. Inaccessible cleavage sites
Nab4p
2. Cleavage complex movement blocked

deleted region also contains the cleavage sites for CP II and CP III. Since cleavage is more permissive in the absence of Nab4p, we reasoned that *cyc1-512* would be competent for cleavage under these relaxed conditions. This was indeed the case, with cleavage giving rise to just two fragments of sizes consistent with cleavage at the remaining CP I site (L. Minvielle-Sebastia, personal communication). This observation was a direct prediction of our model, and in complete agreement with the proposal that Nab4p is necessary during cleavage to stabilize the cleavage complex and enhance specific cleavage events. Recently, the 68 kDa subunit of CF I_m was identified and shown to resemble spliceosomal SR proteins (Rüegsegger et al., 1998). Pre-incubation of CF I_m with RNA precursors enhances the efficiency of cleavage in a manner analogous to the role of SR proteins in pre-mRNA splicing and it was proposed that CF I_m commits the pre-mRNA to cleavage. Though there are only general similarities between the structures of Nab4p and the 68 kDa CF I_m subunit, Nab4p could play a similar role in yeast, whereby specific cleavage enhanced by Nab4p precludes subsequent cleavage of the same transcript.

The effect of Nab4p on polyadenylation has not been studied as extensively as its role in cleavage. However, the *in vivo* data presented here would suggest that its role is limited at best. Neither overexpression of Nab4p or mutations in *NAB4* (including the reconstructed *hrp1-5/nab4-9* allele) has any significant effect on poly(A) tail length, unlike mutations in *RNA14*, *RNA15*, and *PAP1*. In addition, extracts from *nab4* mutants are competent for the polyadenylation reaction at both permissive and non-permissive temperatures. This is in contrast to previous data, which suggested that Hrp1p/Nab4p is required for polyadenylation (Kessler et al., 1997). It is possible that Nab4p also affects

polyadenylation, or the coupling of 3'-end endonucleolytic cleavage to polyadenylation, in a transcript-specific manner, whereas *RNA14*, *RNA15* and *PAP1* are necessary for the polyadenylation of all mRNAs. The requirement for Nab4p during *in vitro* polyadenylation has recently been corroborated (L. Minvielle-Sebastia, personal communication).

It is interesting to note that the major cytoplasmic polyadenylate-tail binding protein Pab1p also co-purifies with CF I, and upon further fractionation, co-purifies with CF IA (Minvielle-Sebastia et al., 1997) or CF IB (Kessler et al., 1997). Pab1p seems to play a role in the control of poly(A) tail length, as strains containing mutations in *PAB1* produce abnormally long poly(A) tails (Sachs and Davis, 1989). Extracts prepared from *pab1* mutants are active for cleavage, but also produce hyperpolyadenylated products. Normal poly(A) tail length can be restored *in vitro* by the addition of recombinant Pab1p (Minvielle-Sebastia et al., 1997 and Kessler et al., 1997). This observation seemed a bit puzzling, since Pab1p has been shown to localize predominantly, if not exclusively to the cytoplasm of cells (see Figure 4). Previous studies have demonstrated that Pab1p is required for translation initiation and also appears to participate in mRNA decapping and deadenylation in the cytoplasm (reviewed by Sachs et al., 1997). However, these recent results suggest that Pab1p resides in the nucleus of cells as well, although undetectable by indirect cellular immunofluorescence techniques. Alternatively, cytoplasmic Pab1p may adventitiously associate with CF I during fractionation. This might explain the differences seen during sub-fractionation of CF I into CF IA and CF IB, which is probably due to variations between biochemical fractionation procedures. If this is the case, then Pab1p may be contributing a non-specific activity to polyadenylation, and the

same may be true of Nab4p, another major poly(A)$^+$ RNA-binding protein. Pab1p is greater than 50% identical to mammalian PAB I, although this mammalian cytoplasmic protein is not known to be involved in nuclear pre-mRNA 3'-end processing. If Pab1p is indeed required for poly(A) tail length control in yeast, then this could explain the apparent lack of a PAB II homologue in yeast.

The Yeast hnRNP A/B-Type Protein Nab4p Has No Apparent Role in Splice Site Selection during Pre-mRNA Splicing

Based again on the similarities between Nab4p and hnRNP A/B proteins, it was reasonable to think that Nab4p may play a role in pre-mRNA splicing. It has been well established that hnRNP proteins, acting antagonistically with SR protein splicing factors, modulate the selection of splice sites, primarily between alternative 5' splice sites, during pre-mRNA splicing (Mayeda and Krainer, 1992). Yet, a direct function in the splicing reaction has not been confirmed for any hnRNPs. This may be complicated by the presence of multiple, highly similar isoforms of hnRNPs, particularly the A/B proteins, in metazoan cells. These related proteins, at least in part, may serve redundant functions and compensate for one another *in vivo*. This could obscure or minimize the effect of mutating or losing a single member of a subgroup of metazoan hnRNPs. Studying the role of a potential splicing factor in yeast is simplified by the apparent lack of alternative splicing patterns in this organism. Any defect in splicing is more likely due to a direct effect on the splicing reaction itself rather than in the choice of endonucleolytic cleavage sites utilized. Also, we know that Nab4p is required for cell viability, indicating that there is not another yeast protein that can supplement the essential function of Nab4p. Yet contrary to our original hypothesis, neither increasing the concentration of Nab4p in

cells nor the loss of Nab4p function had any notable effect of pre-mRNA splicing. There is still the possibility that Nab4p is required for regulated splicing. For example, the intron-containing *MER2* gene is transcribed during both mitosis and meiosis, but is only efficiently spliced during meiosis (Engebrecht et al., 1992). Removal of the *MER2* intron is dependent upon the meiosis-specific Mer1p, a KH motif-containing RNA-binding protein that binds to sequences in the 5' exon and intron of the *MER2* pre-mRNA (Nandabalan and Roeder, 1995). In cases such as this, where the choice is not between multiple splice sites, but rather, whether or not to splice, an hnRNP like Nab4p might serve as a regulatory component. Certainly, the observations with Nab4p do not preclude the possibility that metazoan hnRNP A/B proteins participate in the splicing reaction. It simply indicates that any direct contribution to the splicing reaction by hnRNP A/B-type proteins has not been conserved between yeast and humans, or that such a yeast hnRNP has not been identified.

The Nucleocytoplasmic Shuttling hnRNP Nab4p Is Not Required for mRNA Export

Because of their stable association with pre-mRNA and mRNA *in vivo*, it is not surprising that a number of metazoan hnRNPs have been implicated in a variety of stages during mRNA biogenesis. The same has been true of yeast hnRNPs. Mutational analysis of Nab2p and Npl3p has implicated both of these hnRNPs in mRNA export (Anderson, 1995; Flach et al.; 1994, Singleton et al., 1995; Lee et al., 1996). Several pieces of evidence suggested that Nab4p might also influence mRNA export. The striking similarities between Nab4p and the nucleocytoplasmic shuttling metazoan hnRNP A/B proteins suggested that Nab4p might also be a shuttling hnRNP and have the capacity to

be involved in mRNA export. Indeed, Nab4p is a shuttling hnRNP (Kessler et al., 1997) as is its closest yeast relative, Npl3p (Lee et al., 1996). More recently, Nab2p has been added to the list of yeast hnRNPs that shuttle between the nucleus and cytoplasm of cells (Truant et al., 1998). Nab2p and Nab4p physically interact *in vivo* while *NPL3* and *NAB4/HRP1* have been shown to interact genetically (Henry et al., 1996). *HRP1* was isolated by Henry et al. as a suppressor of an *npl3* mutant, which exhibited a defect in the distribution of poly(A)$^+$ RNA. Together, these results supported a model in which Nab2p and Nab4p, and possibly Npl3p, are components of exported mRNPs. This idea gained further support by the observed block to mRNA export during overexpression of Nab4p. On the other hand, overexpression of Nab4p may block mRNA export, but indirectly, since mutations in *NAB4* do not lead to defects in mRNA export. The role of Nab2p in the export process has not been defined, but if it is essential for mRNA export, then overexpression of Nab4p may sequester Nab2p and inhibit its role in export. To address this possibility, it was determined whether co-overexpression of Nab2p could suppress the toxicity of Nab4p overexpression. Cells containing plasmid-borne alleles of both *NAB2* and *NAB4* expressed from the *GAL1* promoter were unable to grow on galactose, indicating that elevated levels of Nab2p could not alleviate the toxic effect of Nab4p overexpression. However, the maximal production of Nab2p attained by galactose induction was considerably less than that observed for Nab4p, such that Nab4p probably remained in excess of Nab2p. Subsequent genetic studies employing multiple conditional alleles of *NAB4* revealed no obvious effect on mRNA export. Therefore, while Nab4p is quite likely to be a component of at least a subset of exported mRNPs, it is clearly not required for mRNA export from the nucleus.

Since Nab2p has also been shown to influence poly(A) tail length, sequestration of Nab2p by Nab4p might affect pre-mRNA processing at this level as well. Alternatively, the influence of each of these primary nuclear RNA-binding proteins on both 3'-end processing and mRNA export and their putative interaction *in vivo* may point to a possible means by which 3'-end formation and mRNA export are coupled.

Potential for Novel hnRNP Functions in Cytoplasmic mRNA Metabolism

Also unique to Nab4p among other yeast hnRNPs is its ability to accumulate in the cytoplasm of cells during certain growth conditions. Although Npl3p/Nab1p, Nab2p and Nab4p/Hrp1p have all been shown to shuttle between the nucleus and cytoplasm of cells, Nab4p is the only one that accumulates in the cytoplasm, suggesting a cytoplasmic function for this hnRNP. During the induction of mitochondrial biogenesis, approximately half of the Nab4 protein disappears from the nuclei of cells and appears to redistribute in the vicinity of mitochondria in the cytoplasm of cells. The cytoplasmic Nab4 protein is not newly synthesized protein, as cycloheximide does not prevent this accumulation, and there is no change in the overall steady-state concentration of Nab4p in cells (A.M. Krecic and M.S. Swanson, unpublished data). Therefore, it appears that the accumulation of Nab4p in the cytoplasm results from the movement of Nab4p from the nucleus to the cytoplasm, which is not shuttled immediately back into the nucleus. Since Nab4p is an hnRNP, it is possible that Nab4p leaves the nucleus of cells in association with exported mRNPs and remains associated with the mRNA in the cytoplasm. This is true for a number of hnRNPs, including the A/B-type protein hnRNP A1 (Pinol-Roma and Dreyfuss, 1992) and *C. tentans* hrp36 (Visa et al., 1996). In fact,

hrp36 remains associated with its cognate RNA, even in polysomes during ongoing translation. However, there is no evidence regarding the functional significance of this observation.

There is no direct evidence that Nab4p is involved in cytoplasmic mRNA trafficking, but a number of observations suggest that Nab4p may be involved in mRNA localization. First and foremost is the ability of Nab4p to accumulate in the cytoplasm during normal metabolic adjustments to meet the requirements of changing growth conditions. Induction of mitochondrial biogenesis and the heat shock response are means by which budding yeast adjust to such changes in the growth environment (Grivell, 1989; Craig, 1992). Both require induced transcription of a subset of nuclear-encoded genes with a concomitant global repression in the transcription of others. Both of these shifts in gene expression also result in the cytoplasmic accumulation of Nab4p. Induction of mitochondrial biogenesis requires transcriptional activation of many nuclear genes and the proteins encoded by a large number of these genes must be delivered to the mitochondria. Delivery of proteins to mitochondria is mediated, at least in part, by the presence of a mitochondrial targeting signal in these proteins. Another possible mechanism to increase the efficiency of mitochondrial protein targeting would be to localize the corresponding mRNAs in proximity to mitochondria. This could enhance the import of mitochondrial proteins by increasing the local concentration of these proteins and by decreasing the amount of unnecessary translation that may result from the inefficient targeting of mitochondrial proteins to the mitochondria over longer cytoplasmic distances. The observations of mitochondrial-associated polysomes and the enrichment in these polysomes of nuclear-encoded mitochondrial messages supports the

existence of a mRNA localization pathway such as this (Kellems and Butow, 1972; Ades and Butow, 1980). A number of cytoplasmic RNA-binding proteins have been identified that are involved in the localization of mRNAs. A 36 kDa protein that binds to the localization sequence of myelin basic protein mRNA has been identified, and peptide fragments from this protein correspond to hnRNP A2 (Bassell and Singer, 1997). We also know that mRNA localization does occur in *S. cerevisiae*, as demonstrated by the asymmetric distribution of *ASH1* mRNA (Long et al., 1997; Takizawa et al., 1997). Finally, Nab4p interacts with Mgm101p, a protein that is required for maintenance of the mitochondrial genome (Chen et al., 1993). Although the physiological relevance of this interaction has not been explored, it presents a possible link between Nab4p and mitochondrial biogenesis, independent of the Nab4p redistribution results.

Taken together, these observations support the idea that Nab4p might play a role in mRNA trafficking. To test this model it will be important to determine whether Nab4p is associated with mRNAs in the cytoplasm, first by assaying polysomes for the presence of Nab4p followed by UV-crosslinking and subcellular fractionation studies. Immunofluorescence localization of Nab4p coupled with in situ hybridization using probes for a variety of mRNAs would also help determine the specificity of such a pathway. If *NAB4* is required for mRNA localization, then these experiments may be complemented by performing a genetic screen for *nab4* mutants defective in mitochondrial biogenesis.

A second potential point of cytoplasmic regulation of gene expression by Nab4p is the regulation of translation. The functional link between transcription and translational masking suggests that hnRNPs may be important in establishing or

maintaining silent mRNPs. One could predict such a component to be a nucleocytoplasmic shuttling hnRNP that may accumulate in the cytoplasm of cells, remaining associated with the mRNP until the appropriate time for unmasking. To my knowledge, translational silencing has not been documented in yeast. However, Nab4p would be a good candidate for an hnRNP that might influence this phenomenon. Alternatively, hnRNP A/B-type proteins, like Nab4p and hrp36, may stimulate the rate of translation, as hrp36-containing BR mRNA polysomes are translationally active.

Finally, Nab4p is similar in structure and RNA-binding preference to metazoan hnRNP proteins, including the hnRNP D proteins. The first RBD of Nab4p is most similar to one found in AUF1/hnRNP D, suggesting that these two proteins may be similar in their capacity to bind RNA *in vivo*. Analysis of steady-state mRNA levels in *nab4* mutants revealed no global rapid loss of mRNAs at the non-permissive temperature. For example, although levels of *CUP1* mRNA were diminished more rapidly than *TPI1* mRNA levels, which showed little difference from wild type levels, loss of *ACT1* mRNA seemed to occur at an intermediate rate. Loss of these mRNAs in *nab4* mutants could simply be related to epigenetic effects on pre-mRNA processing. To more clearly assay for a direct effect on mRNA stability, mRNA decay rates for a variety of both stable and unstable mRNAs (Parker et al., 1991) would have to be determined and compared between wild type and mutant alleles of *NAB4*.

Conclusions

I report here the independent identification of Nab4p/Hrp1p, an authentic and novel yeast hnRNP, and also demonstrate that it is involved in 3'-end formation of yeast

pre-mRNAs *in vivo*. However, while precursor RNAs with appropriate cleavage and polyadenylation signals can be processed in its absence, Nab4p does play an important role in the correct processing of precursors both *in vitro* and *in vivo*. It is important to note, however, that this was seen for most but not all transcripts tested, suggesting that subsets of pre-mRNAs respond differently during 3'-end formation in conjunction with the concentration of Nab4p. That this could be the case was demonstrated *in vitro*, using a variety of precursors. For example, *CYC1*, which contains relatively strong poly(A) signal sequences, was cleaved efficiently in the absence of Nab4p. In contrast, *GAL7* was cleaved by fractions devoid of Nab4p, but not to the same extent as *CYC1*. This is probably a reflection of the relative strength of polyadenylation signals found in different pre-mRNAs and may reflect a similarity in the role of Nab4p to that proposed for Ref2p in anchoring the cleavage apparatus to the pre-mRNA substrate (Russnak et al., 1995). However, unlike *NAB4*, *REF2* is non-essential and Ref2p does not influence the specificity of cleavage site selection, but rather enhances the utilization of sub-optimal poly(A) sites.

One limitation of the analysis of *nab4* mutants is that all of these mutants have mutations that affect one or both RBD motifs. Targeting mutagenesis outside the RBDs, in regions of the protein that look like domains for protein-protein interactions, may be more fruitful in elucidating other functions of Nab4p. This presumes that the multiple functions for Nab4p involve interactions with discrete sets of proteins, as opposed to all functions of Nab4p involving the interaction of Nab4p with the pre-mRNA/mRNA. Candidates for targeted mutagenesis would include the RGG box motifs and the DR dipeptide repeat in the carboxyl terminus. Mutagenesis of the Nab4p NLS sequence will

certainly be important to continue investigating the localization of Nab4p. Nab4p can also be methylated, presumably at arginine residues within the RGG boxes (Christian Siebel, personal communication) and this post-translational modification may be important in regulating the function of a subset of hnRNPs (Liu and Dreyfuss, 1995; Siebel and Guthrie, 1996; Henry et al., 1996). For example, methylated Npl3p appears to be specific to the nucleus of cells, however the functional significance of this is unknown. Mutations within the RGG boxes that interfere with this post-translational modification may provide important clues regarding the regulation of Nab4p function.

The results presented here clearly predict multiple functions for Nab4p in mRNA biogenesis, which is likely to be true for other hnRNPs. Thus, it would seem that the genetic and biochemical strategies employed thus far have not revealed all of the potential functions of Nab4p. One explanation for this is that I simply have not examined enough of the possible phenotypes that might reveal defects in pre-mRNA processing. However, it is also possible that Nab4p has multiple functions that cannot be uncoupled. Hence, mutations in *NAB4* may affect multiple steps in pre-mRNA processing, to greater or lesser degrees. This highlights an important limitation of *in vitro* assays for pre-mRNA processing and mRNA export, in that they generally measure one activity and not the ability to couple multiple activities. If any process assayed individually does not appear to be significantly affected, then that process may be dismissed as one involving the function of Nab4p. It may in fact be the combination of mild defects at multiple steps in the pathway that leads to lethality in *nab4* mutants. As an abundant primary nuclear pre-mRNA binding proteins, the need for Nab4p and other Nab proteins may be so fundamental to pre-mRNA processing that mutations which cause significant or specific

defects in pre-mRNA processing or mRNA transport may be lethal. This would preclude their isolation by conventional conditional-lethal screens.

This research has had an impact on the way in which we view the regulation of gene expression by hnRNPs. First and most specifically, defining the involvement of Nab4p as an auxiliary specificity factor during pre-mRNA 3'-end processing has revealed new aspects of the mechanism of 3'-end endonucleolytic cleavage. These will be important to incorporate into the current models being tested to elucidate the functions of the various components essential for cleavage, including identification of the endonuclease. Demonstrating that Nab4p influences cleavage site selection *in vivo* provides continued support for the use of the *in vitro* reconstitution system as a reliable representation of pre-mRNA 3'-end processing. Second, the identification of hnRNPs in *S. cerevisiae* suggested that this organism would provide a simplified genetic system for investigating the role of hnRNPs in the regulation of gene expression. This was supported by many subsequent studies, which implicated these proteins in multiple steps during pre-mRNA processing. Defining a specific function for Nab4p in pre-mRNA 3'-end processing illustrates the utility of *S. cerevisiae* in elucidating the contributions of hnRNPs to individual steps during the regulation of gene expression. This work also stresses the importance of keeping an open mind to the possibility that Nab proteins possess multiple functions, some of which may be novel when viewed in the context of current models for the pathways which regulate gene expression. Further, these functions may not be revealed or understood simply by employing techniques which assay one step of pre-mRNA processing uncoupled from others. Finally, the possibility of a cytoplasmic function for Nab4p in yeast is an exciting prospect that may lead to the

discovery of novel cytoplasmic functions for hnRNPs. The cytoplasmic accumulation of Nab4p in *S. cerevisiae* provides a powerful new model for investigating the role of hnRNPs in the cytoplasmic regulation of gene expression.

APPENDIX
YEAST STRAINS, OLIGONUCLEOTIDES AND PLASMIDS

Designation	Genotype/Sequence/Description
Yeast Strains	
BJ926	*MATα/a prb1-1122/prb1-1122 prc1-407/prc1-407 pep4-3/pep4-3 can1/can1 gal2/gal2 his1/HIS1 trp1/TRP1* (Yeast Genetic Stock Center)
S150-2B	*MATa leu2-3; 112 his3Δ200 trp1-289 ura3-52*
LDY133	*MATα ade2 ade3 leu2 trp1 ura3 his3*
YJA213	*MATα leu2 his3Δ200 trp1-289 ura3-52 nab2Δ1::LEU2* [pNAB2.13] (Anderson, 1995)
YJA501	*MATα/a leu2Δ2/leu2Δ2 ura3-52/ura3-52* (Anderson, 1995)
YJA517	YJA501 *NAB2/nab2Δ1::LEU2* [pGAL::nab2] (Anderson, 1995)
YAO403	YJA501 *NAB4/nab4Δ2::LEU2*
YAO404	YAO403 [pNAB4.7]
YAO404-1B	*MATα leu2Δ2 ura3-52 nab4Δ2* [pNAB4.7]
YAO405	YAO404-1B X S150-2B
YAO405-2A	*MATα leu2 ura3-52 trp1-289 his 3Δ200 nab4Δ2::LEU2* [pNAB4.7]
YAO407	YAO403 [pNAB4.10]
YAO408	YAO403 [pRD53]

YAO415	Y190 [pGBTNAB4]
YAO417-1D	*MATa leu2Δ2 ura3-52 nab4Δ2::LEU2* [pNAB4.46]
YAO418-1D	*MATa leu2Δ2 ura3-52 nab4Δ2::LEU2* [pNAB4.47]
YAO420	*MATα leu2 ura3-52 trp1-289 his3Δ200 nab4Δ2::LEU2 TRP1::nab4-1*
YAO421	*MATα ade2 ade3 leu2 trp1 ura3 his4 nab4-1*
YAO428-1B	*MATa leu2Δ2 ura3-52 nab4Δ2::LEU2* [pNAB4.55]
YAO431-1C	*MATa leu2Δ2 ura3-52 nab4Δ2::LEU2* [pNAB4.58]
YAO455	YAO403 [pNAB4.81]
YAO456	YAO403 [pNAB4.81a]
YAO462	YAO403 [pNAB4.88]
YAO467	YAO403 [pNAB4.95]

Oligonucleotides

MSS015	5'-CGTAATAATGGCTACCATCCATATAATAGGGGA ATTCTATAAAAAAAAAGTGTAG-3'
MSS177	5'-CTAATACACGTACATGT-3'
MSS181	5'-TGGTAGCCATTATTACG-3
MSS192	5'-GCAGTATACACAAGTCT-3'
MSS194	5'-GGTATACTCAATAAACA-3'
MSS197	5'-CCCGGATCCTTTTTACCTATTATATG-3'
MSS198	5'-CCCGAATTCGCAATGAGCTCTGACGA-3'
MSS221	5'-CCCGGATCCAGTCTGAGAAAATAGAG-3'
MSS222	5'-CCCAAGCTTGACAGCTGTACTTC-3'

MSS513	5'-GCTGATGTAGTAGAAGATCC-3'
MSS5110	5'-GGGGAGCTCTATGGGTAAAGGAGAAG-3'
MSS5111	5'-GGGGAGCTCCCCTTGTATAGTTCATCC-3'
MSS5123	5'-CCCGAATTCCCCTATTATATGGATGGTAG-3'
MSS5124	5'-CCCGAGCTCCCCAATATATGGATGGTAG-3'
MSS565	5'-CACTAGAAGGTTTTTCAAAAGATGAGAAACCGA ACCCTCTAG-3'
MSS564	5'-CTAGAGGGTTCGGTTTCTCATCTTTTGAAAAACC TTCTAGTG
MSS511	5'-CCCGAATTCGCCTCTATTTTCC-3'
MSS514	5'-GGATCCGGTATCAAAACG-3'
MSS553	5'GGGGGATCCTGTTTCTAATCTCTGCTTTTGTGCG CGT-3'
MSS569	5'-GCAGGTCGATATCATGTAATTAGTTATG-3'
MSS574	5'-GGGGGATCCCGCGTGTACGCATGTAACATTATA CTG

Plasmids

pNAB2.13 2.46 kb PvuI/NsiI fragment from *nab2-13*, blunt-ended and subcloned into SmaI-cut pRS316 (Sikorski and Hieter, 1989)

pGAL::nab2 BglII/SalI fragment from pNAB2-3B6 (Anderson et al., 1993b) subcloned into BamHI/SalI-cut pRD53 [(EcoRI/BamHI fragment containing the *GAL1,10* promoter region subcloned into SpeI/BamHI-cut pRS316) gift of R.J. Deshaies, University of California, San Francisco, CA]

pRNP4.0 2.3 kb EcoRI fragment of RNP2-13.1 (starting at nucleotide 96 of the *NAB4* gene) isolated by expression screening a λgt11 genomic DNA library and subcloned into EcoRI-cut pSP72 (Promega, Madison, WI)

pRNP4.6	EcoRI/HindIII fragment from pRNP4.0 subcloned into EcoRI/HindIII-cut pSP72
pRNPex4.7	EcoRI/SalI fragment of pRNP4.0 subcloned into EcoRI/SalI-cut pGEX-4T-1 (Promega, Madison, WI)
pMAL4.2	EcoRI/BamHI-cut PCR product amplified using MSS214 and MSS215, encoding the amino-terminal RGG box of Nab4p (amino acids 316-405) fused in-frame to the maltose-binding protein by subcloning into EcoRI/BamHI-cut pMAL-c2 (New England Biolabs, Beverly, MA)
pMAL4.3	EcoRI/BamHI-cut PCR product amplified using MSS216 and MSS217, encoding the carboxy-terminal RGG box of Nab4p (amino acids 445-534) fused in-frame to the maltose-binding protein by subcloning into EcoRI/BamHI-cut pMAL-c2
pNAB4p.0	*NAB4* coding region (nt –3 to 1608), produced by PCR using MSS197 and MSS198 and inserted into EcoRI/BamHI-cut pSP72
pGBTNAB4	1.6 kb EcoRI/BamHI fragment of pNAB4p.0 subcloned into EcoRI/BamHI-cut pGBT9 (Clontech)
pNAB4.0	YCp50 clone containing entire *NAB4* gene
pNAB4.7	2.7 kb BamHI/SalI fragment from pNAB4.0 subcloned into BamHI/SalI-cut pRS316
pNAB4.7A	750 bp BstEII/HindIII fragment of secondary PCR product generated using MSS222 and MSS243 with combined 5' and 3' *NAB4* primary PCR products generated using MSS243 and MSS181 or MSS222 and MSS015, respectively, cloned into BstEII/HindIII-cut pNAB4.7
pNAB4.8	2.7 kb BamHI/SalI fragment of pNAB4.0 subcloned into pRS316
pNAB4.8A	pNAB4.8 cut linearized with SstI, blunt-ended and religated to eliminate the SacI site in the polylinker
pNAB4.9	2.2 kb XhoI/SalI fragment YEp13 containing *LEU2*, blunt-ended and subcloned into StyI-cut and blunt-ended pRNP4.6
pNAB4.10	*NAB4* coding region (nt –34 to 1787), produced by PCR using MSS221 and MSS222, inserted into BamHI/HindIII-cut pRD53

pNAB4.13	KpnI/BamHI-cut PCR product from pMAL4.2 generated using MSS58 and MSS215 subcloned into KpnI/BamHI-cut pSP72
pNAB4.14	KpnI/BamHI-cut PCR product from pMAL4.3 generated using MSS58 and MSS217 subcloned into KpnI/BamHI-cut pSP72
pNAB4.36	2.7 kb BamHI/SalI fragment containing *nab4-1* subcloned into BamHI/SalI-cut pRS304 (Sikorski and Hieter, 1989)
pNAB4.42	pNAB4.36 cut with NotI and BamHI, blunt-ended, and religated
pNAB4.46	2.7 kb BamHI/SalI fragment of pNAB4.0 subcloned into BamHI/SalI-cut YCp50
pNAB4.47	2.7 kb BamHI/SalI fragment containing *nab4-1* subcloned into BamHI/SalI-cut YCp50
pNAB4.55	2.7 kb BamHI/SalI fragment containing *nab4-4* subcloned into BamHI/SalI-cut YCp50
pNAB4.58	2.7 kb BamHI/SalI fragment containing *nab4-7* subcloned into BamHI/SalI-cut YCp50
pNAB4.76	725 bp SstI-cut GFP PCR product generated from pS65T-C1 (Clontech, Palo Alto, CA) using MSS5110 and MSS5111 subcloned in frame with *NAB4* into SstI-cut pNAB4.46
pNAB4.80	2.4 kb BamHI/BstEII fragment of pNAB4.76 subcloned into BamHI/BstEII-cut pNAB4.7A
pNAB4.80a	AatII/EcoRI-cut pNAB4.80, blunt-ended and religated (258 bp deletion)
pNAB4.80b	BstEII/EcoRI-cut pNAB4.80, blunt-ended and religated (561 bp deletion)
pNAB4.80c	ClaI/EcoRI-cut pNAB4..80, blunt-ended and religated (792 bp deletion)
pNAB4.80d	EcoNI/EcoRI-cut pNAB4..80, blunt-ended and religated (1020 bp deletion)
pNAB4.80e	NruI/EcoRI-cut pNAB4..80, blunt-ended and religated (1075 bp deletion)

pNAB4.80f	Eco47III/EcoRI-cut pNAB4..80, blunt-ended and religated (1305 bp deletion)
pNAB4.80g	KpnI-cut pNAB4..80, blunt-ended and religated (1500 bp deletion)
pNAB4.81	0.7 kb SstI fragment of GFP PCR product amplified from pS65T-C1 (Clontech, Palo Alto, CA) using MSS5110 and MSS5111, subcloned into SstI-cut pNAB4.8A
pNAB4.81a	3.1 kb BamHI/SalI fragment of pNAB4.80a subcloned into BamHI/SalI-cut YCp50
pNAB4.81b	2.8 kb BamHI/SalI fragment of pNAB4.80a subcloned into BamHI/SalI-cut YCp50
pNAB4.81c	2.6 kb BamHI/SalI fragment of pNAB4.80c subcloned into BamHI/SalI-cut YCp50
pNAB4.81d	2.4 kb BamHI/SalI fragment of pNAB4.80d subcloned into BamHI/SalI-cut YCp50
pNAB4.81e	2.3 kb BamHI/SalI fragment of pNAB4.80e subcloned into BamHI/SalI-cut YCp50
pNAB4.81f	2.1 kb BamHI/SalI fragment of pNAB4.80f subcloned into BamHI/SalI-cut YCp50
pNAB4.81g	1.5 kb BamHI/SalI fragment of pNAB4.80g subcloned into BamHI/SalI-cut YCp50
pNAB4.88	700 bp SstI GFP fragment from pNAB4.80 plus 171bp SstI/EcoRI-cut PCR product generated from pNAB4.80 using MSS5123 and MSS5124 subcloned into SstI/EcoRI-cut pNAB4.81
pNAB4.95	1.1 kb Sst1/BstEII fragment of secondary PCR product generated using MSS198 and MSS222 with combined 5' and 3' *nab4* primary PCR products generated using MSS198 and MSS565 or MSS222 and MSS564, respectively, cloned into SstI/BstEII-cut pNAB4.46
pACT1.1	EcoRI/BamHI-cut *ACT1* fragment including 5' and 3' UTR sequences, PCR amplified using MSS511 and MSS515, cloned into EcoRI/BamHI-cut pSP72

REFERENCES

Ades, I.Z., and R.A. Butow. 1980. The products of mitochondria-bound cytoplasmic polysomes in yeast. *J. Biol. Chem.* **255**:9918-9924.

Ainger, K., D. Avossa, F. Morgan, S.J. Hill, C. Barry, E. Barbarese, and J.H. Carson. 1993. Transport and localization of exogenous MBP mRNA microinjected into oligodendrocytes. *J. Cell Biol.* **123**:431-441.

Aitchison, J.D., G. Blobel, and M.P. Rout. 1996. Kap104p: a karyopherin involved in the nuclear transport of messenger RNA binding proteins. *Science* **274**:624-627.

Akada, R., J. Yamamoto, and I. Yamashita. 1997. Screening and identification of yeast sequences that cause growth inhibition when overexpressed. *Mol. Gen. Genet.* **254**:267-274.

Altschul, S.F., W. Gish, W. Miller, E.W. Myers, and D.J. Lipman. 1990. Basic alignment search tool. *J. Mol. Biol.* **215**:403-410.

Alzhanova-Ericsson, A.T., X. Sun, N. Visa, E. Kiseleva, T. Wurtz, and B. Daneholt. 1996. A protein of the SR family of splicing factors binds extensively to exonic Balbiani ring pre-mRNA and accompanies the RNA from the gene to the nuclear pore. *Genes Dev.* **10**:2881-2893.

Amberg, D.C., A.L. Goldstein, and C.N. Cole. 1992. Isolation and characterization of *RAT1*: and essential gene of *Saccharomyces cerevisiae* required for the efficient nucleocytoplasmic trafficking of mRNA. *Genes Dev.* **6**:1173-1189.

Amero, S.A., G. Raychaudhuri, C.L. Cass, W.J. van Venrooij, W.J. Habets, A.R. Krainer, and A.L. Beyer. 1992. Independent deposition of heterogenous nuclear ribonucleoproteins and small nuclear ribonucleoprotein particles at sites of transcription. *Proc. Natl. Acad. Sci. USA* **89**:8409-8413.

Amrani, N., M. Minet, M. Le Gouar, F. Lacroute, and F. Wyers. 1997a. Yeast Pab1p interacts with Rna15 and participates in the control of the poly(A) tail length in vitro. *Mol. Cell Biol.* **17**:3694-3701.

Amrani, N., M. Minet, F. Wyers, M.-E. Dufour, L.P. Aggerbech, and F. Lacroute. 1997b. *PCF11* encodes a third protein component of yeast cleavage and polyadenylation factor I. *Mol. Cell. Biol.* **17**:1102-1109.

Anderson, J.T. 1995. Characterization of nuclear polyadenylated RNA-binding proteins from *Saccharomyces cerevisiae*. Ph.D. Dissertation, University of Florida.

Anderson, J.T., M.R, Paddy, and M.S. Swanson. 1993a. PUB1 is a major nuclear and cytoplasmic polyadenylated RNA-binding protein in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **13**:6102-6113.

Anderson, J.T., S.M. Wilson, K.V. Datar, and M.S. Swanson. 1993b. NAB2: a yeast nuclear polyadenylated RNA-binding protein essential for cell viability. *Mol. Cell. Biol.* **13**:2730-2741.

Antoniou, M. 1995. Clinical defects in pre-mRNA processing. *In* A.I. Lamond (ed.), Pre-mRNA Processing. R.G. Landes Company, Austin, Texas, p. 187-201.

Bai, C., and P.P. Tolias. 1998. Drosophila clipper/CPSF 30K is a post-transcriptionally regulated nuclear protein that binds RNA containing GC clusters. *Nuc. Acids Res.* **26**:1597-1604.

Ballantyne, S., A. Bilger, J. Aström, A. Virtanen, and M. Wickens. 1995. Poly(A) polymerases in the nucleus and cytoplasm of frog oocytes: dynamic changes during oocyte maturation and early development. *RNA* **1**:64-78.

Bassell, G., and R.H. Singer. 1997. mRNA and cytoskeletal filaments. *Curr. Opin. Cell Biol.* **9**:109-115.

Biamonti, G., M. Ruggiu, S. Saccone, G. Della Valle, and S. Riva. 1994. Two homologous genes, originated by duplication, encode the human hnRNP proteins A2 and A1. *Nuc. Acids Res.* **22**:1996-2002.

Bilger, A., C.A. Fox, E. Wahle, and M. Wickens. 1994. Nuclear polyadenylation factors recognize cytoplasmic polyadenylation elements. *Genes Dev.* **8**:1106-1116.

Boelens, W.C., C. Dargemont, and I.W. Mattaj. 1995. Export mRNA through the nuclear pore complex. *In* A.I. Lamond (ed.), Pre-mRNA Processing. R.G. Landes Company, Austin, Texas, p. 173-186.

Bossie, M.A., C. DeHoratius, G. Barcelo, and P. Silver. 1992. A mutant protein with similarity to RNA binding proteins interferes with nuclear import in yeast. Mol. Biol. *Cell* **3**:875-893.

Bouvet, P., and A.P. Wolffe. 1994. A role for transcription and FRGY2 in masking maternal mRNA within Xenopus oocytes. *Cell* **77**:931-941.

Burd, C.G., and G. Dreyfuss. 1994. RNA binding specificity of hnRNP A1: significance of hnRNP A1 high-affinity binding sites in pre-mRNA splicing. *EMBO J.* **13**:1197-1204.

Burd, C.G., M.S. Swanson, M. Gorlach, and G. Dreyfuss. 1989. Primary structures of the heterogeneous nuclear ribonucleoprotein A2, B1, and C2 proteins: a diversity of RNA binding proteins is generated by small peptide inserts. *Proc. Natl. Acad. Sci. USA* **86**:9788-9792.

Cáceres, J.F., S. Stamm, D.H. Helfman, A.R. Krainer. 1994. Regulation of alternative splicing in vivo by overexpression of antagonistic splicing factors. *Science* **265**:1706-1709.

Chambers, A. 1996. Highly conserved toxicity of *Saccharomyces cerevisiae* Rap1p. *Mol. Microbiol.* **22**:449-458.

Chanfreau, G., S.M. Noble, and C. Guthrie. 1996. Essential yeast protein with unexpected similarity to subunits of mammalian cleavage and polyadenylation specificity factor (CPSF). *Science* **274**:1511-1514.

Chase, J.W., K.R. Williams. 1986. Single-stranded DNA binding proteins required for DNA replication. *Ann. Rev. Biochem.* **55**:103-136.

Chen, F.C., C. MacDonald, and J. Wilusz. 1995. Cleavage site determinants in the mammalian polyadenylation signal. *Nuc. Acids Res.* **23**:2614-2620.

Chen, X.-J., M.-X. Guan, and G.D. Clark-Walker. 1993. *MGM101*, a nuclear gene involved in maintenance of the mitochondrial genome in *Saccharomyces cerevisiae*. *Nuc. Acids Res.* **21**:3473-3477.

Choi, Y.D., P.J. Grabowski, P.A. Sharp, and G. Dreyfuss. 1986. Heterogeneous nuclear ribonucleoproteins: role in RNA splicing. *Science* **231**:1534-1539.

Colgan, D.F., and J.L. Manley. 1997. Mechanism and regulation of mRNA polyadenylation. *Genes Dev.* **11**:2755-2766.

Coppola, J.A., A.S. Field, and D.S. Luse. 1983. Promoter-proximal pausing by RNA polymerase II *in vitro*: transcripts shorter than 20 nucleotides are not capped. *Proc. Natl. Acad. Sci. USA* **80**:1251-1255.

Craig, E.A. 1992. The heat-shock response of *Saccharomyces cerevisiae*. *In* E.W. Jones, J.R. Pringle, and J.R. Broach, (eds.), The Molecular and Cellular Biology of the Yeast Saccharomyces. Cold Spring Harbor Laboratory Press, New York, p. 501-537.

Curtis, D. R. Lehmann, and P.D. Zamore. 1995. Translational regulation in development. *Cell* **81**:171-178.

Decker, C.J., and R. Parker. 1995. Diversity of cytoplasmic functions for the 3' untranslated region of eukaryotic transcripts. *Curr. Opin. Cell Biol.* **7**:386-392.

DeMaria, C.T., and G. Brewer. 1996. AUF1 binding affinity to A+U-rich elements correlated with rapid mRNA degradation. *J. Biol. Chem.* **271**:12179-12184.

Dohmen, R.J., P. Wu, and A. Varshavsky. 1994. Heat inducible degron: a method for constructing temperature-sensitive mutants. *Science* **263**:1273-1276.

Drazinic, C.M., J.B. Smerage, M.C. Lopez, and H.V. Baker. 1996. Activation mechanism of the multifunctional transcription factor repressor-activator protein 1 (Rap1p). *Mol. Cell Biol.* **16**:3187-3196.

Dreyfuss, G., M.J. Matunis, S. Piñol-Roma, and C.G. Burd. 1993. hnRNP proteins and the biogenesis of mRNA. *Ann. Rev. Biochem.* **62**:289-321.

Eckner, R. W. Ellmeier, and M.L. Birnstiel. 1991. Mature mRNA 3' end formation stimulates RNA export from the nucleus. *EMBO J.* **10**:3513-3522.

Edwalds-Gilbert, G., K.L. Veraldi, and C. Milcarek. 1997. Alternative poly(A) site selection in complex transcription units: means to an end? *Nuc. Acids Res.* **25**:2547-2561.

Ehrenmann, K., L. Long, B.J. Wagner, and G. Brewer. 1994. Characterization of cDNAs encoding the murine A+U-rich RNA-binding protein AUF1. *Gene* **149**:315-319.

Eng, F.J., and J.R. Warner. 1991. Structural basis for the regulation of splicing of a yeast messenger RNA. *Cell* **65**:797-804.

Engebrecht, J.A., K. Voelkel-Meiman, and G.S. Roeder. 1992. Meiosis-specific RNA splicing in yeast. *Cell* **66**:1257-1268.

Ernst, R.K., M. Bray, D. Rekosh, and M.-L. Hammarskjöld. 1997. A structured retroviral RNA element that mediated nucleocytoplasmic export of intron-containing RNA. *Mol. Cell Biol.* **17**:135-144.

Espinet C., M.A. de la Torre, M. Aldea, and E. Herrero. 1995. An efficient method to isolate yeast genes causing overexpression-mediated growth arrest. *Yeast* **11**:25-32.

Fields, S., and R. Sternglanz. 1994. The two-hybrid system: an assay for protein-protein interactions. *Trends Genet.* **10**:286-292.

Flach, J., M. Bossie, J. Vogel, A. Corbett, T. Jinks, D.A. Willins, and P.A. Silver. 1994. A yeast RNA-binding protein shuttles between the nucleus and cytoplasm. *Mol. Cell. Biol.* **14**:8399-8407.

Forrester W., F. Stutz, M. Rosbash, and M. Wickens. 1992. Defects in mRNA 3'-end formation, transcription initiation, and mRNA transport associated with the yeast mutation prp20: possible coupling of mRNA processing and chromatin structure. *Genes Dev.* **6**:1914-1926.

Fox, C.A., M.D. Sheets, and M.P. Wickens. 1989. Poly(A) addition during maturation of frog oocytes: distinct nuclear and cytoplasmic activities and regulation by the sequence UUUUUAU. *Genes Dev.* **3**:2151-2162.

Furth, P., W. Choe, J. Rex, J.C. Byrne, and C. Baker. 1994. Sequences homologous to 5' splice sites are required for the inhibitory activity of papillomavirus late 3' untranslated regions. *Mol. Cell. Biol.* **14**:5278-5289.

Gamberi, C., G. Contreas, M.G. Romanelli, and C. Morandi. 1994. Analysis of the yeast *NSR1* gene and protein domain comparison between Nsr1 and human hnRNP type A1. *Gene* **148**:59-66.

Ghetti, A., S. Pinol-Roma, W.M. Michael, C. Morandi, and G. Dreyfuss. 1992. hnRNP I, the polypyrimidine tract-binding protein: distinct nuclear localization and association with hnRNAs. *Nuc. Acids Res.* **20**:3671-3678.

Gil, A., P.A. Sharp, S.F. Jamison, and M.A. Garcia-Blanco. 1991. Characterization of cDNAs encoding the polypyrimidine tract-binding protein. *Genes Dev.* **5**:1224-1236.

Görlach, M., M. Wittekind, R.A. Beckman, L. Mueller, and G. Dreyfuss. 1992. Interaction of the RNA-binding domain of the hnRNP C proteins with RNA. *EMBO J.* **11**:3289-3295.

Görlich, D., and I.W. Mattaj. 1996. Nucleocytoplasmic transport. *Science* **271**:1513-1518.

Grivell, L.A. 1989. Nucleo-mitohondrial interactions in yeast mitochondrial biogenesis. *Eur. J. Biochem.* **182**:477-493.

Guthrie, C., and G.R. Fink. 1991. Guide to Yeast Genetics and Molecular Biology. Academic Press, San Diego, CA.

Guo, Z., P. Russo, D.F. Yun, J.S. Butler, and F. Sherman. 1995. Redundant 3'-end formation signals for the yeast *CYC1* mRNA. *Proc. Natl. Acad. Sci. USA* **92**:4211-4214.

Guo, Z., and F. Sherman. 1995. 3'-end-forming signals of yeast mRNA. *Mol. Cell. Biol.* **15**:5983-5990.

Guo, Z., and F. Sherman. 1996. 3'-End forming signals of yeast mRNA. *Trends Biochem.Sci.* **21**:477-481.

Hake, L.E., and J.D. Richter. CPEB is a specificity factor that mediates cytoplasmic polyadenylation during Xenopus oocyte maturation. *Cell* **79**:617-627.

Harper, J.W., G.R. Adami, N. Wei, K. Keyomarski, and S.J. Elledge. 1993. The p21 Cdk-interacting protein Cip1 is a potent inhibitor or G1 cyclin-dependent kinases. *Cell* **75**:805-816.

Heim, R., A.B. Cubitt, and R.Y. Tsien. 1995. Improved green fluorescence. *Nature* **373**:663-664.

Henry, M., C.Z. Borland, M. Bossie, and P.A. Silver. 1996. Potential RNA binding proteins in *Saccharomyces cerevisiae* identified as suppressors of temperature-sensitive mutations in *NPL3*. *Genetics* **142**:103-115.

Henry, M.F., and P.A. Silver. 1996. A novel methyltransferase (Hmt1p) modifies poly(A)+-RNA-binding proteins. *Mol. Cell. Biol.* **16**:3668-3678.

Huang, Y., and G.C. Carmichael. 1996. Role of polyadenylation in nucleocytoplasmic transport of mRNA. *Mol. Cell Biol.* **16**:1534-1542.

Humphrey, T., P. Sadhale, T. Platt, and N. Proudfoot. 1991. Homologous mRNA 3' end formation in fission and budding yeast. *EMBO J.* **10**:3503-3511.

Izaurralde, E., A. Jarmolowski, C. Beisei, I.W. Mattaj, G. Dreyfuss, and U. Fischer. 1997. A role for the M9 transport signal of hnRNP A1 in mRNA nuclear export. *J. Cell Biol.* **137**:27-35.

Izaurralde, E., and I.W. Mattaj. 1995. RNA export. *Cell* **81**:153-159.

Jarmolowski, A., W.C. Boelens, E. Izaurralde, and I.W. Mattaj. 1994. Nuclear export of different classes of RNA is mediated by specific factors. *J. Cell Biol.* **124**:627-635.

Jenny, A., L. Minvielle-Sebastia, P.J. Preker, and W. Keller. 1996. Sequence similarity between the 73-kilodalton protein of mammalian CPSF and a subunit of yeast polyadenylation factor I. *Science* **274**:1514-1517.

Jessen, T.H., C. Oubridge, C.H. Teo, C. Pritchard, and K. Nagai. 1991. Identification of molecular contacts between the U1 A small nuclear ribonucleoprotein and U1 RNA. *EMBO J.* **10**:3447-3456.

Kellems, R.E., and R.A. Butow. 1972. Cytoplasmic-type 80S ribosomes associated with yeast mitochondria. *J. Biol. Chem.* **247**:8043-8050.

Keller. W. 1995. 3'-End cleavage and polyadenylation of nuclear messenger RNA precursors. *In* A.I. Lamond (ed.), Pre-mRNA Processing. R.G. Landes Company, Austin, Texas, p. 113-134.

Keller, W., and L. Minvielle-Sebastia. 1997. A comparison of mammalian and yeast pre-mRNA 3'-end processing. *Curr. Opin. Cell Biol.* **9**:329-336.

Kessler, M.M., M.F. Henry, E. Shen, J. Zhao, S. Gross, P.A. Silver, and C.L. Moore. 1997. Hrp1, a sequence-specific RNA-binding protein that shuttles between the nucleus and the cytoplasm, is required for mRNA 3'-end formation in yeast. *Genes Dev.* **11**:2545-2556.

Kessler, M.M., Zhao, J., and C.L. Moore. 1996. Purification of the *Saccharomyces cerevisiae* cleavage/polyadenylation factor I. *J. Biol. Chem.* **271**:27167-27175.

Kiledjian, M., and G. Dreyfuss. 1992. Primary structure and RNA binding activity of the hnRNP U protein: binding RNA through RGG box. *EMBO J.* **7**:265502664.

Kiledjian, M., C.T. DeMaria, G. Brewer, and K. Novick. 1997. Identification of AUF1 (heterogeneous nuclear ribonucleoprotein D) as a component of the alpha-globin mRNA stability complex. *Mol. Cell. Biol.* **17**:4870-4876.

Knowles, R.B., J.H. Sabry, M.A. Martone, M. Ellisman, G.J. Bassell, and K.S. Kosik. 1996. Translocation of RNA granules in living neurons. *J. Neurosci.* **16**:7812-7820.

Kramer, A. 1995. The biochemisty of pre-mRNA splicing. *In* A.I. Lamond (ed.), Pre-mRNA Processing. R.G. Landes Company, Austin, Texas, p. 35-64.

Krawczak, M., J. Reiss, and D.N. Cooper. 1992. The mutational spectrum of single base-pair substitutions in mRNA splice junctions of human genes: causes and consequences. *Hum. Genet.* **90**:41-54.

Lee, F.J., and J. Moss. 1993. An RNA-binding protein gene (*RBP1*) of *Saccharomyces cerevisiae* encodes a putative glucose-repressible protein containing two RNA recognition motifs. *J. Biol. Chem.* **268**:15080-15087.

Lee, M.S., M. Henry, and P.A. Silver. 1996. A protein that shuttles between the nucleus and the cytoplasm is an important mediator of RNA export. *Genes Dev.* **10**:1233-1246.

Lewis, J.D., and E. Izaurralde. 1997. The role of the cap structure in RNA processing and nuclear export. *Eur. J. Biochem.* **247**:461-469.

Lin, J.J., and V.A. Zakian. 1994. Isolation and characterization of two *Saccharomyces cerevisiae* genes that encode proteins that bind to $(TG_{1\text{-}3})_n$ single strand telomeric DNA *in vitro*. *Nuc. Acids Res.* **22**:4906-4913.

Lingner, J., J. Kellerman, and W. Keller. 1991. Cloning and expression of the essential gene for poly(A) polymerase from yeast. *Nature* **354**:496-498.

Liu, C., X. Mao, and A.J. Lustig. 1994. Mutational analysis defines a C-terminal tail domain of RAP1 essential for telomeric silencing in *Saccharomyces cerevisiae*. *Genetics* **138**:1025-1040.

Liu, H., J. Krizek, and A. Bretscher. 1992. Construction of a *GAL1*-regulated yeast cDNA expression library and its application to the identification of genes whose overexpression causes lethality in yeast. *Genetics* **132**:665-673.

Liu, Q., and G. Dreyfuss. 1995. In vivo and in vitro arginine methylation of RNA-binding proteins. *Mol. Cell. Biol.* **15**:2800-2808.

Long, R.M., R.H. Singer, X. Meng, I. Gonzalez, K. Nasmyth, and R.-P. Jansen. Mating type switching in yeast controlled by asymmetric localization of *ASH1* mRNA. *Science* **277**:383-387.

Longtine, M.S., N.M. Wilson, M.E. Petracek, and J. Berman. 1989. A yeast telomeric binding activity binds to two related telomere sequence motifs and is indistinguishable from RAP1. *Curr. Genet.* **6**:225-239.

Lou, H., Y. Yang, G.J. Cote, and S.M. Berget. 1996. An intron enhancer recognized by splicing factors activates polyadenylation. *Genes Dev.* **10**:208-219.

Lutz, C.S., and J.C. Alwine. 1994. Direct interaction of the U1 snRNP-A protein with the upstream efficiency element of the SV40 late polyadenylation signal. *Genes Dev.* 8:576-586.

MacDonald, C.C., J. Wilusz, and T. Shenk. 1994. The 64-kilodalton subunit of the CstF polyadenylation factor binds to pre-mRNAs downstream of the cleavage site and influences cleavage site selection. *Mol. Cell. Biol.* **14**:6647-6654.

Mandart, E., and R. Parker. 1995. Effects of mutations in the *Saccharomyces cerevisiae RNA14*, *RNA15*, and *PAP1* genes on polyadenylation in vivo. *Mol. Cell. Biol.* **15**:6979-6986.

Mann, K.P., E.A. Weiss, and J.R. Nevins. 1993. Alternative poly(A) site utilization during adenovirus infection coincides with a decrease in the activity of a poly(A) site processing factor. *Mol. Cell. Biol.* **13**:2411-2419.

Martin, G., and W. Keller. 1996. Mutational analysis of mammalian poly(A) polymerase identifies a region for primer binding and a catalytic domain, homologous to the family X polymerases, and t other nucleotidyltransferases. *EMBO J.* **15**:2593-2603.

Matunis, E.L., M.J. Matunis, and G. Dreyfuss. 1993. Association of individual hnRNP proteins and snRNPs with nascent transcripts. *J. Cell Biol.* **121**:219-228.

Mayeda, A., and A.R. Krainer. 1992. Regulation of alternative pre-mRNA splicing by hnRNP A1 and splicing factor SF2. *Cell* **68**:365-375.

Mehlin, H., and B. Daneholt. 1993. The Balbiani ring particle: a model for the assembly and export of RNPs from the nucleus? *Trends Cell Biol.* **3**:443-447.

Mehlin, H., B. Daneholt, and U. Skoglund. 1992. Translocation of a specific premessenger ribonucleoprotein particle through the nuclear pore studied with electron microscope tomography. *Cell* **69**:605-613.

Michael, W.M., M. Choi, and G. Dreyfuss. 1995. A nuclear export signal in hnRNP A1: signal-mediated, temperature-dependent nuclear protein export pathway.*Cell* **83**:415-422.

Michael, W.M., P.S. Eder, and G. Dreyfuss. 1997. The K nuclear shuttling domain: a novel signal for nuclear import and nuclear export in the hnRNP K protein. *EMBO J.* **16**:3587-3598.

Minvielle-Sebastia, L., P.J. Preker, and W. Keller. 1994. RNA14 and RNA15 as components of a yeast pre-mRNA 3'-end processing factor. *Science* **266**:1702-1705.

Minvielle-Sebastia, L., P.J. Preker, T. Wiederkehr, Y. Strahm, and W. Keller. 1997. The major yeast poly(A)-binding protein is associated with cleavage factor IA and functions in premessenger RNA 3'-end formation. *Proc. Natl. Acad. Sci. USA* **94**:7897-7902.

Minvielle-Sebastia, L., B. Winsor, N. Bonneaud, and F. Lacroute. 1991. Mutations in the yeast *RNA14* and *RNA15* genes result in an abnormal mRNA decay rate; sequence analysis reveals an RNA-binding domain in the RNA15 protein. *Mol. Cell. Biol.* **11**:3075-3087.

Moore, M.S., and G. Blobel. 1994. A G protein involved in nucleocytoplasmic transport: the role of Ran. *Trends Biochem. Sci.* **19**:211-216.

Muhlrad, D., R. Hunter, and R. Parker. 1992. A rapid method for localized mutagenesis of yeast genes. *Yeast* **8**:79-82.

Munroe, S.H., and X. Dong. 1992. Heterogeneous nuclear ribonucleoprotein A1 catalyzes RNA•RNA annealing. *Proc. Natl. Acad. Sci. USA* **89**: 895-899.

Murphy, R., and S.R. Wente. 1996. An RNA-export mediator with an essential nuclear export signal. *Nature* **383**:357-360.

Nakamura, M., H. Okano, J.A. Blendy, and C. Montell. 1994. Musashi, a neural RNA-binding protein required for *Drosophila* adult external organ development. *Neuron* **13**:67-81.

Nakielny, S., and G. Dreyfuss. 1996. The hnRNP C proteins contain a nuclear retention sequence that can override nuclear export signals. *J. Cell Biol.* **134**:1365-1373.

Nakielny, S., and G. Dreyfuss. 1997. Nuclear export of proteins and RNAs. *Curr. Opin. Cell Biol. 9*:420-429.

Nakielny, S., M.C. Siomi, H. Siomi, W.M. Michael, V.W. Pollard, and G. Dreyfuss. 1996. *Exp. Cell Res.* **229**:261-266.

Nandabalan, K., and G.S. Roeder. 1995. Binding of a cell-type-specific RNA splicing factor to its target regulatory sequence. *Mol. Cell. Biol.* **15**:1953-1960.

Nasmyth, K., and R.-P. Jansen. 1997. The cytoskeleton in mRNA localization and cell differentiation. *Curr. Opin. Cell Biol.* **9**:396-400.

Nemeth, A., S. Krause, D. Blank, A. Jenny, P. Jenö, A. Lustig, and E. Wahle. 1995. Isolation of genomic and cDNA clones encoding bovine poly(A) binding protein II. *Nuc. Acids Res.* **23**:4034-4011.

O'Hare, K. 1995. mRNA 3'ends in focus. *Trends Genet.* **11**:255-257.

Patton, J.G., S.A. Mayer, P. Tempst, and B. Nadal-Ginard. 1991. Characterization and molecular cloning of polypyrimidine tract-binding protein: a component of a complex necessary for pre-mRNA splicing. *Genes Dev.* **5**:1237-1251.

Parker, R., D. Herrick, S. Peltz, and A. Jacobson. 1991. Measurement of mRNA decay rates in Saccharomyces cerevisiae. *Methods Enzymol.* **194**:415-423.

Petitjean, A., N. Bonneaud, and F. Lacroute. 1995. The *duplicated Saccharomyces cerevisiae* gene *SSM1* encodes a eukaryotic homolog of the eubacterial and archaebacterial L1 ribosomal proteins. *Mol. Cell. Biol.* **15**:5071-5081.

Piñol-Roma, S., and G. Dreyfuss. 1992. Shuttling of pre-mRNA binding proteins between nucleus and cytoplasm. *Nature* **355**:730-732.

Pollard, V.W., W.M. Michael, S. Nakielny, M.C. Siomi, F. Wang, and G. Dreyfuss. 1996. A novel receptor-mediated nuclear protein import pathway. *Cell* **86**:985-994.

Pontius, B.W., and Berg, P. 1990. Renaturation of complementary DNA strands mediated by purified mammalian heterogeneous nuclear ribonucleoprotein A1 protein: Implications for a mechanism for rapid molecular assembly. *Proc. Natl. Acad. Sci. USA.* **87**:8403-8407.

Portman, D.S., and G. Dreyfuss. 1994. RNA annealing activities in HeLa nuclei. *EMBO J.* **13**:213-221.

Prasher, D.C., V.K. Eckenrode, W.W. Ward, F.G. Prendergast, and M..J. Cormier. 1992. Primary structure of the *Aequorea victoria* green fluorescent protein. *Gene* **111**:229-233.

Preker, P.J., J. Lingner, L. Minvielle-Sebastia, and W. Keller. 1995. The *FIP1* gene encodes a component of a yeast pre-mRNA polyadenylation factor that directly interacts with poly(A) polymerase. *Cell* **81**:379-389.

Preker, P.J., M. Ohnacker, L. Minvielle-Sebastia, and W. Keller. 1997. A multisubunit 3' end processing factor from yeast containing poly(A) polymerase and homlogues of the subunits of mammalian cleavage and polyadenylation specificity factor. *EMBO J.* **16**:4727-4737.

Richter, J.D. 1995. Dynamics of poly(A) addition and removal during development. *In* M. Mathews, J. Hershey, and N. Sonenbert (eds.), Translational Control. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Rose, M.D., F. Winston, and P. Hieter. 1990. Methods in Yeast Genetics. A Laboratory Course Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Rüegsegger, U., K. Beyer, and W. Keller. 1996. Purification and characterization of human cleavage factor I_m involved in 3'end processing of messenger RNA precursors. *J. Biol. Chem.* **271**:6107-6113.

Rüegsegger, U., D. Blank, and W. Keller. 1998. Human pre-mRNA cleavage factor I_m is related to spliceosomal SR proteins and can be reconstituted from recombinant subunits. *Mol. Cell* **1**:243-253.

Russell, I.D., and D. Tollervey. 1992. NOP3 is an essential yeast protein which is required for pre-rRNA processing. *J. Cell Biol.* **119**:737-747.

Russell, I., and D. Tollervey. 1995. Yeast Nop3p has structural and functional similarities to mammalian pre-mRNA binding proteins. *Eur. J. Cell Biol.* **66**:293-301.

Russnak, R., K.W. Nehrke, and T. Platt. 1995. *REF2* encodes an RNA-binding protein directly involved in yeast mRNA 3'-end formation. *Mol. Cell. Biol.* **15**:1689-1697.

Rymond, B.C., and M. Rosbash. 1992. Yeast pre-mRNA splicing. *In* E.W. Jones, J.R. Pringle, and J.R. Broach (eds.), The Molecular and Cellular Biology of the Yeast Saccharomyces. Cold Spring Harbor Laboratory Press, New York, p. 143-192.

Sachs, A.B., M.W. Bond, and R.D. Kornberg. 1986. A single gene from yeast for both the nuclear and cytoplasmic polyadenylate-binding proteins: domain structure and expression. *Cell* **45**:827-835.

Sachs, A.B., and R.W. Davis. 1989. The poly(A) binding protein is required for poly(A) shortening and 60S ribosomal subunit-dependent translation initiation. *Cell* **58**:857-867.

St. Johnston, D. 1995. The intracellular localization of messenger RNAs. *Cell* **81**:161-170.

Sambrook, J., E. Fritsch, and T. Maniatis. 1989. Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Sazer, S. 1996. The search for the primary function of the Ran GTPase continues. *Trends Cell Biol.* **6**:81-85.

Schneider, J.C., and L. Guarente. 1991. Vectors for expression of cloned genes in yeast: regulation, overproduction, and underproduction. *Methods Enzymol.* **194:**373-388.

Shatkin, A.J. 1976. Capping of eukaryotic mRNAs. *Cell* **9**:645-654.

Siebel, C.W., and C. Guthrie. 1996. The essential yeast RNA binding protein Npl3p is methylated. *Proc. Natl. Acad. Sci. USA* **93**:13641-13646.

Sierakowska, H., W. Szer, P.J. Furdon, and R. Kole. 1986. Antibodies to hnRNP core proteins inhibit *in vitro* splicing of human beta-globin pre-mRNA. *Nuc. Acids Res.* **14**:5241-5254.

Sikorski, R.S., and J.D. Boeke. 1991. In vitro mutagenesis and plasmid shuffling: from cloned gene to mutant yeast. *Methods Enzymol.* **194**:302-318.

Sikorski, R.S., and P. Hieter. 1989. A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in *Saccharomyces cerevisiae*. *Genetics* **122**:19-27.

Singleton, D.R., S. Chen, M. Hitomi, C. Kumagai, and A.M. Tartakoff. 1995. A yeast protein that bidirectionally affects nucleocytoplasmic transport. *J. Cell Sci.* **108**:265-272.

Siomi, H., and G. Dreyfuss. 1995. A nuclear localization domain in the hnRNP A1 protein. *J. Cell Biol.* **129**:552-560.

Snyder, M., S. Elledge, D. Sweeter, R.A.Young, and R.W. Davis. 1987. lgt11: gene isolation with antibody probes and applications. *Methods Enzymol.* **154**:107-128.

Strathern, J., and D.R. Higgins. 1991. Recovery of plasmids from yeast into Eschericia coli: Shuttle vectors. *Methods. Enzymol.* **194**:319-329.

Stumpf, G., and H. Domdey. 1996. Dependence of yeast pre-mRNA 3'-end processing on CFT1: a sequence homolog of the mammalian AAUAAA binding factor. *Science* **274**:1517-1520.

Stutz, F., M. Neville, and M. Rosbash. 1995. Identification of a novel nuclear pore-associated protein as a functional target of the HIV-1 Rev protein in yeast. *Cell* **82**:495-506.

Swanson, M.S. 1995. Functions of nuclear pre-mRNA/mRNA binding proteins. *In* A.I. Lamond (ed.), Pre-mRNA Processing. R.G. Landes Company, Austin, Texas, p. 17-33.

Takagaki, Y., C.C. MacDonald, T. Shenk, and J.L. Manley. 1992. The human 64-kDa polyadenylation factor contains a ribonucleoprotein-type RNA binding domain and unusual auxiliary motifs. *Proc. Natl. Acad. Sci. USA* **89**:1403-1407.

Takagaki, Y., R.L Seipelt, M.L. Peterson, and J.L. Manley. 1996. The polyadenylation factor CstF-64 regulates alternative processing of IgM heavy chain pre-mRNA during B-cell differentiation. *Cell* **87**:941-952.

Takizawa, P.A., A. Sil, J.R. Swedlow, I. Herskowitz, and R.D. Vale. 1997. Actin-dependent localization of an RNA encoding a cell-fate determinant in yeast. *Nature* **389**:90-93.

Theissen, H., M. Etzerodt, R. Reuter, C. Schneider, F. Lottspeich, P. Argos, R. Lührmann, and L. Philipson. 1986. Cloning of the human cDNA for the U1 RNA-associated 70K protein. *EMBO J.* **5**:3209-3217.

Truant, R., R.A. Fridell, R.E. Benson, H. Bogerd, and B.R. Cullen. 1998. Identification and functional characterization of a novel nuclear localization signal present in the yeast Nab2 poly(A)$^+$ RNA binding protein. *Mol. Cell. Biol.* **18**:1449-1458.

Varnum, S.M., C.A. Hurney, and W.M. Wormington. 1992. Maturation-specific deadenylation in Xenopus oocytes requires nuclear and cytoplasmic factors. *Dev. Biol.* **153**:283-290.

Visa, N., A.T. Alzhanova-Ericsson, X. Sun, E. Kiseleva, B. Björkroth, T. Wurtz, and B. Daneholt. 1996a. A pre-mRNA-binding protein accompanies the RNA from the gene through the nuclear pores and into polysomes. *Cell* **84**:253-264.

Visa, N., E. Izaurralde, J. Ferreira, B. Daneholt, and I.W. Mattaj. 1996b. A nuclear cap binding complex binds Balbiani ring pre-mRNA co-trancsriptionally and accompanies the ribonucleoprotein particle during nuclear export. *J. Cell Biol.* **133**:5-14.

Wahle, E. 1995. Poly(A) tail length control in caused by termination of processive synthesis. *J. Biol. Chem.* **270**:2800-2808.

Wahle, E., and W. Keller. 1992. The biochemistry of 3'-end cleavage and polyadenylation of messenger RNA precursors. *Ann. Rev. Biochem.* **61**:419-440.

Wahle, E., and W. Keller. 1996. The biochemistry of polyadenylation. *Trends Biochem. Sci.* **21**:247-250.

West, A.H., D.J. Clark, J. Martin, W. Neupert, R.- Ulrich Hartl, and A.L. Horwich. 1992. Two related genes encoding extremely hydrophobic proteins suppress a lethal mutation in the yeast mitochondrial processing enhancing protein. *J. Biol. Chem.* **267**:24625-24633.

Wilhelm, J.E., and R.D. Vale. 1993. RNA on the move: the mRNA localization pathway. *J. Cell Biol.* **123**:269-274.

Wilson, S.M. 1996. Functional analysis of a heterogeneous nuclear ribonucleoprotein in vivo. Ph.D. Dissertation, University of Florida.

Wilson, S.M., K.V. Datar, M.R. Paddy, J.R. Swedlow, and M.S. Swanson. 1994. Characterization of nuclear polyadeylated RNA-binding proteins. *J. Cell Biol.* **127**:1173-1184.

Wilusz, J., and T. Shenk. 1990. A uridylate tract mediates efficient heterogeneous nuclear ribonucleoprotein C protein-RNA cross-linking and functionally substitutes for the downstream element of the polyadenylation signal. *Mol. Cell. Biol.* **10**:6397-6407.

Wu, L., P.J. Good, and J.D. Richter. 1997. The 36-kilodalton embryonic-type cytoplasmic polyadenylation element-binding protein in *Xenopus laevis* is ElrA, a member of the ELAV family of RNA-binding proteins. *Mol. Cell Biol.* **17**:6402-6409.

Yaffe, M.P. 1991. Analysis of mitochondrial function and assembly. *Methods Enzymol.* **194**:627-643.

Yan, C., and T. Melese. 1993. Multiple regions of *NSR1* are sufficient for accumulation of a fusion protein within the nucleolus. *J. Cell Biol.* **123**:1081-1091.

Yue, Z., E. Maldonado, R. Pillutla, H. Cho, D. Reinberg, and A.J. Shatkin. 1997. Mammalian capping enzyme complements mutant *Saccharomyces cerevisiae* lacking mRNA guanylyltransferase and selectively binds the elongating form of RNA polymerase II. *Proc. Natl. Acad. Sci. USA* **94**:12898-12903.

Zaret, K.S., and F. Sherman. 1982. DNA sequence required for efficient transcription termination in yeast. *Cell* **28**:563-573.

Zhao, J., M.M. Kessler, and C.L. Moore. 1997. Cleavage factor II of *Saccharomyces cerevisiae* contains homologues to subunits of the mammalian cleavage/polyadenylation specificity factor and exhibits sequence-specific, ATP-dependent interaction with precursor RNA. *J. Biol. Chem.* **272**:10831-10838.

BIOGRAPHICAL SKETCH

The author was born Annette Marie Oberdorf, in Lewisburg, Pennsylvania, on June 3, 1968. She is the second of three daughters raised by Larry E. and Bonita L. Oberdorf. She began her undergraduate education at Juniata College in Huntingdon, Pennsylvania, in the fall of 1986. She explored a career in science as a summer intern, first at the Sigfried and Janet Weis Center for Research of the Geisinger Clinic in Danville, Pennsylvania, and later as a Whitaker scholar at the Milton S. Hershey Medical Center in Hershey, Pennsylvania. After graduating cum laude with a Bachelor of Science degree in biology, she continued her pursuit of a career in science as a research technician at the University of Massachusetts Medical Center in Worcester, Massachusetts.

With the desire to pursue a career as a college professor, Ms. Oberdorf began her graduate studies at the University of Florida in the fall of 1992 and embarked on her pre-doctoral training under the guidance of Dr. Maurice S. Swanson. During graduate school, she befriended and later married Michael R. Krecic. She received her doctoral degree in the spring of 1998.

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Maurice S. Swanson, Chair
Associate Professor of Molecular Genetics and Microbiology

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Alfred S. Lewin
Professor of Molecular Genetics and Microbiology

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Sue A. Moyer
Professor of Molecular Genetics and Microbiology

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

William W. Hauswirth
Eminent Scholar of Molecular Genetics and Microbiology

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Thomas W. O'Brien
Professor of Biochemistry and Molecular Biology

This dissertation was submitted to the Graduate Faculty of the College of Medicine and to the Graduate School and was accepted as partial fulfillment of the requirements for the degree of Doctor of Philosophy.

May 1998

Dean, College of Medicine

Dean, Graduate School